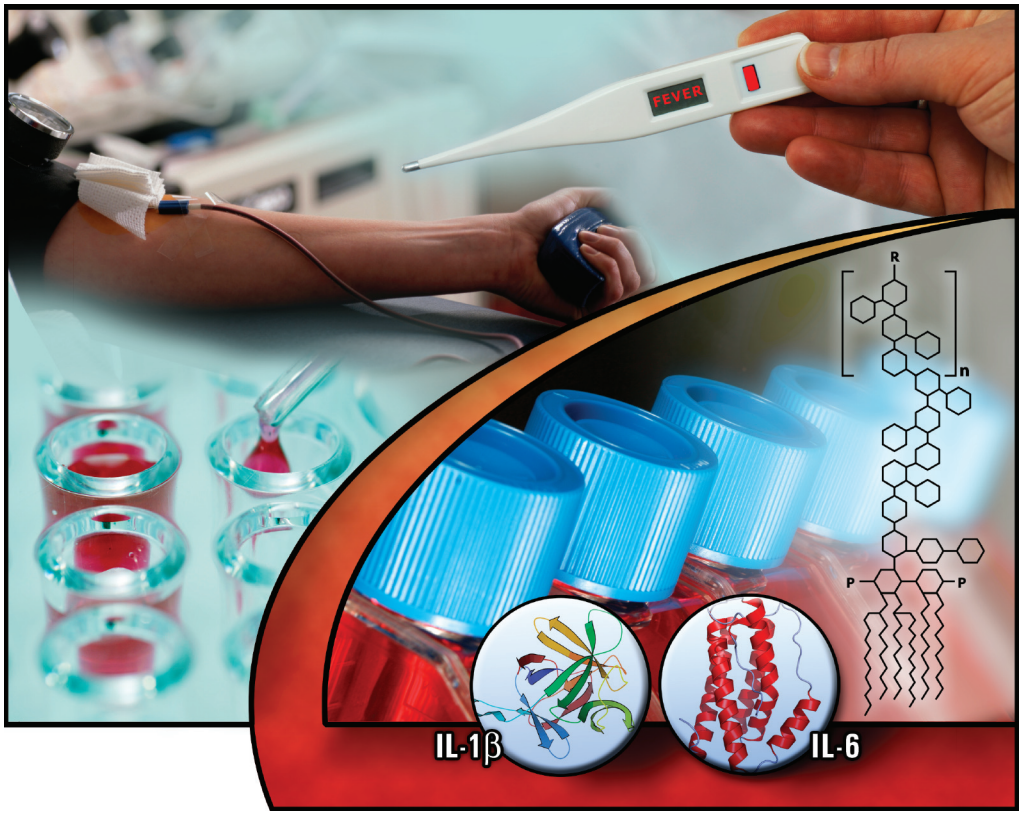
NIH Publication Number 08-6391



ICCVAM BACKGROUND REVIEW DOCUMENT

Validation Status of Five *In Vitro* Test Methods Proposed for Assessing Potential Pyrogenicity of Pharmaceuticals and Other Products

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)

National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

National Institute of Environmental Health Sciences
National Institutes of Health
U. S. Public Health Service
Department of Health and Human Services

About the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)
and
The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

In 1997, the National Institute of Environmental Health Sciences (NIEHS), one of the National Institutes of Health (NIH), established ICCVAM to:

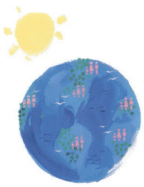
- Coordinate interagency technical reviews of new and revised toxicological test methods, including alternative test methods that reduce, refine, or replace the use of animals
- Coordinate cross-agency issues relating to validation, acceptance, and national and international harmonization of new, modified, and alternative toxicological test methods

On December 19, 2000, the ICCVAM Authorization Act (Public Law 106-545, 42 U.S.C. 285*l*-3) established ICCVAM as a permanent interagency committee of NIEHS under NICEATM.

ICCVAM is comprised of representatives from 15 U.S. Federal regulatory and research agencies that use, generate, or disseminate toxicological information. ICCVAM conducts technical evaluations of new, revised, and alternative methods with regulatory applicability. ICCVAM promotes the scientific validation and regulatory acceptance of toxicological test methods that more accurately assess the safety or hazards of chemicals and products and that refine (decrease or eliminate pain and distress), reduce, and/or replace animal use. NICEATM administers ICCVAM and provides scientific and operational support for ICCVAM-related activities. More information about ICCVAM and NICEATM can be found on the NICEATM-ICCVAM web site (http://iccvam.niehs.nih.gov) or obtained by contacting NICEATM (telephone: [919] 541-2384, e-mail: niceatm@niehs.nih.gov).

The following Federal regulatory and research agencies are ICCVAM members:

- Consumer Product Safety Commission
- Department of Agriculture
- Department of Defense
- Department of Energy
- Department of Health and Human Services
  - Centers for Disease Control and Prevention
    - Agency for Toxic Substances and Disease Registry
    - National Institute of Occupational Safety and Health
  - Food and Drug Administration
  - National Institutes of Health
    - Office of the Director
    - National Cancer Institute
    - National Institute of Environmental Health Sciences
    - National Library of Medicine
- Department of the Interior
- Department of Labor
  - Occupational Safety and Health Administration
- Department of Transportation
- Environmental Protection Agency



The NICEATM-ICCVAM graphic symbolizes the important role of new and alternative toxicological methods in protecting and advancing the health of people, animals, and our environment.

ICCVAM Background Review Document:

Validation Status of Five *In Vitro* Test Methods Proposed for Assessing Potential Pyrogenicity of Pharmaceuticals and Other Products

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)

National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

National Institute of Environmental Health Sciences
National Institutes of Health
U.S. Public Health Service
Department of Health and Human Services

May 2008
NIH Publication No. 08-6391

National Toxicology Program
P.O. Box 12233
Research Triangle Park, NC 27709

This document is available electronically at:
http://iccvam.niehs.nih.gov/methods/pyrogen/pyrogen.htm

Table of Contents

Page Number

List of Tables ..... **vi**
List of Abbreviations and Acronyms ..... **vii**
Interagency Coordinating Committee on the Validation of Alternative Methods: Agency Representatives ..... **x**
Acknowledgements ..... **xi**
Preface ..... **xiii**
Executive Summary ..... **xvii**

1.0 Introduction and Rationale for the Proposed Use of *In Vitro* Pyrogen Test Methods **1-1**
1.1 Introduction 1-1
1.1.1 Historical Background of *In Vitro* Pyrogen Test Methods and the Rationale for Their Development 1-1
1.1.2 Peer Reviews of *In Vitro* Pyrogen Test Method Validation Studies 1-3
1.2 Regulatory Rationale and Applicability 1-3
1.2.1 Current Regulatory Testing Requirements and ICCVAM Prioritization Criteria 1-3
1.2.2 Intended Uses of the Proposed *In Vitro* Pyrogen Test Methods 1-7
1.2.3 Similarities and Differences in the Endpoints Measured by the Proposed Test Methods and the *In Vivo* Reference Test Method 1-7
1.2.4 Use of the Proposed Test Methods in an Overall Strategy of Hazard or Safety Assessment 1-7
1.3 Scientific Basis for the *In Vitro* Pyrogen Test Methods 1-8
1.3.1 Purpose and Mechanistic Basis of the *In Vitro* Pyrogen Test Methods 1-8
1.3.2 Similarities and Differences of Modes of Action Between the *In Vitro* Pyrogen Test Methods and the Fever Response in Humans and/or Rabbits 1-8
1.3.3 Range of Substances Amenable to the *In Vitro* Pyrogen Test Methods and Limits of These Methods 1-9
1.4 Validation of the *In Vitro* Pyrogen Test Methods 1-9
1.5 Search Strategies and Selection of Citations for the ICCVAM BRD 1-10

2.0 *In Vitro* Pyrogen Test Method Protocol Components **2-1**
2.1 Overview of How the *In Vitro* Pyrogen Test Methods Are Conducted 2-1
2.2 Description and Rationale for the Test Method Components for Proposed Standardized Protocols 2-1
2.2.1 Methods Used to Analyze the Data, Including Methods to Analyze for Interference with the Assay 2-7
2.2.2 Decision Criteria and the Basis for the Prediction Model Used to Identify a Pyrogenic Substance 2-7

2.2.3 Information and Data to be Included in the Study Report and Availability of Standard Forms for Data Collection and Submission 2-8
2.3 Basis for Selection of the Test Method Systems 2-9
2.4 Proprietary Components 2-10
2.5 Number of Replicates 2-11
2.5.1 Number of Donors 2-11
2.5.2 Number of Assay Replicates 2-11
2.6 Modifications to the Test Method Protocols Based on ECVAM Validation Study Results 2-11
2.7 Differences Between Comparable Validated Test Methods with Established Performance Standards 2-11

3.0 Substances Used for the Validation of *In Vitro* Pyrogen Test Methods 3-1
3.1 Rationale for the Substances or Products Selected for Testing 3-1
3.2 Number of Substances 3-1
3.3 Identification and Description of Substances Tested 3-1
3.4 Sample Coding Procedure 3-2
3.5 Rationale for the Selection of the Recommended Reference Substances 3-3

4.0 *In Vivo* Reference Data for the Assessment of Test Method Accuracy 4-1
4.1 Description of the Protocol Used to Generate *In Vivo* Data 4-1
4.1.1 The Rabbit Pyrogen Test (RPT) 4-1
4.1.2 Current *In Vivo* Pyrogen Test Method Protocols 4-5
4.2 Reference Data Used to Assess *In Vitro* Test Method Accuracy 4-6
4.3 Availability of Original Records for the *In Vivo* Reference Data 4-6
4.4 *In Vivo* Data Quality 4-6
4.5 Availability and Use of Toxicity Information from the Species of Interest 4-7
4.6 Information on the Accuracy and Reliability of the *In Vivo* Test Method 4-7

5.0 Test Method Data and Results 5-1
5.1 Test Method Protocol 5-1
5.2 Availability of Copies of Original Data Used to Evaluate Test Method Performance 5-1
5.3 Description of the Statistical Approaches Used to Evaluate the Resulting Data 5-1
5.4 Summary of Results 5-2
5.5 Use of Coded Chemicals and Compliance with GLP Guidelines 5-3
5.6 Lot-to-Lot Consistency of Test Substances 5-3
5.7 Availability of Data for External Audit 5-3

6.0 Relevance of the *In Vitro* Pyrogen Test Methods 6-1
6.1 Accuracy of *In Vitro* Pyrogen Test Methods 6-1
6.1.1 Relevance of the Cryo WB/IL-1β Test Method 6-2
6.1.2 Relevance of the MM6/IL-6 Test Method 6-2

6.1.3 Relevance of the PBMC/IL-6 Test Method....6-3
6.1.3.1 Relevance of the PBMC/IL-6 Method When Using Cryo PBMCs....6-3
6.1.4 Relevance of the WB/IL-6 Test Method....6-3
6.1.5 Relevance of the WB/IL-1β Test Method....6-3
6.1.5.1 Relevance of the WB/IL-1β Test Method When Using 96-Well Plates....6-3
6.2 Summary of the Performance Statistics for *In Vitro* Pyrogen Test Methods....6-4
6.2.1 Discordant Results....6-4
6.2.2 Strengths and Limitations of *In Vitro* Pyrogen Test Methods....6-4

7.0 Reliability of the *In Vitro* Pyrogen Test Methods....7-1
7.1 Selection Rationale for the Substances Used to Evaluate the Reliability of *In Vitro* Pyrogen Test Methods....7-1
7.2 Analysis of Intralaboratory Repeatability and Reproducibility....7-1
7.2.1 Intralaboratory Repeatability....7-2
7.2.2 Intralaboratory Reproducibility....7-3
7.2.3 Interlaboratory Reproducibility....7-7
7.3 Historical Positive and Negative Control Data....7-8

8.0 Test Method Data Quality....8-1
8.1 Adherence to National and International GLP Guidelines....8-1
8.2 Data Quality Audits....8-1
8.3 Impact of Deviations from GLP Guidelines....8-1
8.4 Availability of Laboratory Notebooks or Other Records....8-1
8.5 Need for Data Quality....8-1

9.0 Other Scientific Reports and Reviews....9-1
9.1 Summaries of *In Vitro* Pyrogen Test Methods and Data from Published and Unpublished Studies....9-1
9.1.1 Andrade et al. (2003)....9-1
9.1.2 Bleeker et al. (1994)....9-4
9.1.3 Carlin and Viitanen (2003)....9-4
9.1.4 Carlin and Viitanen (2005)....9-7
9.1.5 Daneshian et al. (2006)....9-7
9.1.6 Eperon et al. (1996, 1997)....9-8
9.1.7 Marth and Kleinhappl (2002)....9-10
9.1.8 Martis et al. (2005)....9-11
9.1.9 Pool et al. (1998)....9-11
9.1.10 Taktak et al. (1991)....9-13
9.2 Conclusions from Scientific Literature Based on Independent Peer-Reviewed Reports and/or Reviews....9-13
9.2.1 De Groote et al. (1992)....9-14
9.2.2 Fennrich et al. (1999)....9-14
9.2.3 Hansen and Christensen (1990)....9-15
9.2.4 Hartung and Wendel (1996)....9-15
9.2.5 Moesby et al. (1999)....9-15

9.2.6 Nakagawa et al. (2002) 9-16
9.2.7 Pool et al. (1999) 9-16
9.2.8 Poole et al. (2003) 9-16
9.2.9 Schindler et al. (2004) 9-17

10.0 Animal Welfare Considerations (Refinement, Reduction, and Replacement) 10-1
10.1 How the Five *In Vitro* Test Methods Will Refine, Reduce, or Replace Animal Use 10-1
10.2 Requirement for the Use of Animals 10-2
10.2.1 Rationale for the Use of Animals 10-2

11.0 Practical Considerations 11-1
11.1 Transferability of the *In Vitro* Pyrogen Test Methods 11-1
11.1.1 Facilities and Major Fixed Equipment 11-1
11.1.2 General Availability of Other Necessary Equipment and Supplies 11-1
11.2 Personnel Training Considerations 11-1
11.2.1 Required Training and Expertise Needed to Conduct the *In Vitro* Pyrogen Test Methods 11-2
11.3 Cost Considerations 11-2
11.4 Time Considerations 11-2

12.0 References 12-1

13.0 Glossary 13-1

Appendix A ECVAM BRDs and Standard Operating Procedures A-1
A1 The Human Whole Blood (WB)/Interleukin (IL)-1β *In Vitro* Pyrogen Test A-3
A2 The Human WB/IL-1β *In Vitro* Pyrogen Test: Application of Cryopreserved (Cryo) Human WB A-133
A3 The Human WB/IL-6 *In Vitro* Pyrogen Test A-237
A4 The Human Peripheral Blood Mononuclear Cell (PBMC)/IL-6 *In Vitro* Pyrogen Test A-335
A5 The Monocytoid Cell Line Mono Mac 6 (MM6)/IL-6 *In Vitro* Pyrogen Test A-459

Appendix B ECVAM Response to ICCVAM Questions B-1

Appendix C Additional Information Requested by the Panel C-1
C1 ESAC Statement on the Validity of *In Vitro* Pyrogen Tests C-3
C2 Press Release: "Fewer Tests on Animals and Safer Drugs: New EU Tests Save 200,000 Rabbits per Year" C-9
C3 ECVAM Replies to Questions of ICCVAM Pyrogenicity Peer Review Panel C-15
C4 Rationale for the Selection of the 10 Substances Tested in the Validation/Catch-Up Validation Study of *In Vitro* Assays for Pyrogen Testing C-21
C5 Comparison and Validation of Novel Pyrogen Tests Based on the

Human Fever Reaction: Trial Data Report..... C-25
C6 List of Drugs for the Catch-Up Validation Study..... C-41
C7 Analytical Procedure to Identify and Eliminate Outlying Observations C-45

List of Tables

		Page Number
Table 1*	Accuracy of *In Vitro* Pyrogen Test Methods	xix
Table 1-1	Summary of U.S. and European Legislation and Statutory Protocol Requirements for Pyrogenicity Testing	1-5
Table 2-1	*In Vitro* Pyrogen Test Method Components	2-3
Table 2-2	Prediction Model Used for *In Vitro* Pyrogen Test Methods	2-8
Table 3-1	Parenteral Drugs Used in the Validation Studies for Determining Test Method Accuracy	3-2
Table 3-2	Parenteral Drugs Used in the Validation Studies for Determining Test Method Reproducibility	3-2
Table 4-1	Test Guidelines for the Rabbit Pyrogen Test	4-3
Table 4-2	Decision Criteria for Determining a Pyrogenic Response in the Rabbit Pyrogen Test	4-6
Table 6-1*	Accuracy of *In Vitro* Pyrogen Test Methods	6-2
Table 6-2	Predictivity of *In Vitro* Pyrogen Test Method for Each Endotoxin Spike Concentration	6-5
Table 7-1	Intralaboratory Repeatability Assessed with Saline Spiked with WHO-LPS 94/580	7-3
Table 7-2	Intralaboratory Reproducibility of *In Vitro* Pyrogen Test Methods	7-5
Table 7-3	Interlaboratory Reproducibility of *In Vitro* Pyrogen Test Methods	7-7
Table 7-4	Interlaboratory Reproducibility of *In Vitro* Pyrogen Test Methods	7-8
Table 9-1	Results of Pyrogen Testing of Pharmaceutical/Biological Products in the Human PBMC Assay, the BET, and the RPT	9-2
Table 9-2	Results of Pyrogen Testing of Pharmaceutical/Biological Products by the Human WB Culture Assay, the BET, and the RPT	9-3
Table 9-3	IL-6 Production from WB after Exposure to Endotoxin or Five Infanrix® Vaccines	9-5
Table 9-4	IL-6 Production by MM6 Cells after Exposure to Endotoxin or Five Infanrix® Vaccines	9-6
Table 9-5	Pyrogenic Activity of Blood Preparations for Parenteral Use	9-9
Table 9-6	Pyrogenic Activity of Vaccine Preparations	9-10
Table 9-7	Comparison of the WB Test, BET, and the RPT for Detecting Pyrogens in Production Batches of Biological Products	9-12
Table 9-8	Results of Pyrogen Testing of Batches of Therapeutic HSA Using the MM6/IL-6, BET, and RPT	9-13
Table 9-9	Comparison of the Application Spectra of the RPT, the BET, and the Human WB Assay (PyroCheck®)	9-14
Table 11-1	Cost Estimates for the RPT and *In Vitro* Pyrogen Tests	11-3

*Tables 1 and 6-1 are identical. Table 1 provides supporting information for the Executive Summary, while Table 6-1 provides the information in Section 6.1, where test method accuracy is discussed in greater detail.

List of Abbreviations and Acronyms

AWIPT	Adsorb, Wash, *In Vitro* Pyrogen Test
BET	Bacterial Endotoxin Test
BRD	Background Review Document
CBER	Center for Biologics Evaluation and Research
CCAC	Canadian Council on Animal Care
CDER	Center for Drug Evaluation and Research
CDRH	Center for Devices and Radiological Health
CEC	Commission of the European Communities
CFR	Code of Federal Regulations
CLB	Central Laboratory for the Blood Transfusion Service
Cryo	Cryopreserved
CSF	Colony Stimulating Factor
CV	Coefficient of variation
CVM	Center for Veterinary Medicine
DMSO	Dimethyl sulfoxide
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen (German Collection of Microorganisms and Cell Cultures)
ECVAM	European Centre for the Validation of Alternative Methods
EDQM	European Directorate for the Quality of Medicines
EEC	European Economic Community
ELC	Endotoxin limit concentration
ELISA	Enzyme-linked immunosorbent assay
EMEA	European Medicines Agency
EP	European Pharmacopeia
EPA	U.S. Environmental Protection Agency
ESAC	ECVAM Scientific Advisory Committee
EU	European Union
EU/mL	Endotoxin units/mL
FDA	U.S. Food and Drug Administration
Fn	Fibronectin
FR	*Federal Register*
GLP	Good Laboratory Practice
GM-CSF	Granulocyte-Macrophage-Colony Stimulating Factor
Hb	Hemoglobin
hGH	Human growth hormone
HSA	Human serum albumin
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
IFN-γ	Interferon-γ
IgG	Immunoglobulin G
IL	Interleukin
i.m.	Intramuscular
ISO	International Standards Organization
IU	International units

i.v.	Intravenous
JP	Japanese Pharmacopeia
LAL	*Limulus* Amebocyte Lysate
LPS	Lipopolysaccharide
LTA	Lipoteichoic acid
MEM	Minimum essential medium
MM6	Mono Mac 6
MVD	Maximum valid dilution
NA	Not applicable
NC	Negative control
n.c.	Not calculated
ND	Not done
n.d.	Not detectable
NF	National Formulary
NI	Not included
NICEATM	National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods
NPC	Negative product control
OECD	Organisation for Economic Co-operation and Development
OD	Optical density
PBMC	Peripheral Blood Mononuclear Cell
PC	Positive control
PEI	Paul Ehrlich Institut
PFS	Pyrogen-free saline
PG	Peptidoglycan
PHA	Phytohemagglutinin
PPC	Positive product control
PWG	Pyrogenicity Working Group
QA	Quality assurance
rec	Recombinant
RPT	Rabbit Pyrogen Test
SD	Standard deviation
SOP	Standard operating procedure
TBE	Tick-borne encephalitis
THP-1	Acute monocyte leukemia cell line
TLR	Toll-like Receptor
TNF-α	Tumor Necrosis Factor-α
U	Units
U.K.	United Kingdom
U.S.	United States
U.S.C.	United States Code
USDA	U.S. Department of Agriculture
USP	U.S. Pharmacopeia
USPTO	United States Patent and Trademark Office
UV	Ultraviolet
WB	Whole blood

WHO	World Health Organization
x *g*	Times gravity

Interagency Coordinating Committee on the Validation of Alternative Methods: Agency Representatives

Agency for Toxic Substances and Disease Registry
★Moiz Mumtaz, Ph.D.

Consumer Product Safety Commission
★ Marilyn L. Wind, Ph.D. **(Chair)**
◊ Kristina Hatlelid, Ph.D.
Joanna Matheson, Ph.D.

Department of Agriculture
★ Jodie Kulpa-Eddy, D.V.M. **(Vice-Chair)**
◊ Elizabeth Goldentyer, D.V.M.

Department of Defense
★ Robert E. Foster, Ph.D.
◊ Patty Decot
Peter J. Schultheiss, D.V.M., D.A.C.L.A.M.
Harry Salem, Ph.D.

Department of Energy
★ Michael Kuperberg, Ph.D.
◊ Marvin Stodolsky, Ph.D.

Department of the Interior
★ Barnett A. Rattner, Ph.D.
◊ Sarah Gerould, Ph.D.

Department of Transportation
★ George Cushmac, Ph.D.
◊ Steve Hwang, Ph.D.

Environmental Protection Agency
Office of Science Coordination and Policy
★ Karen Hamernik, Ph.D.
Office of Research and Development
◊ Julian Preston, Ph.D.
Suzanne McMaster, Ph.D.
Office of Pesticides Programs
Amy Rispin, Ph.D.
Deborah McCall
OECD Test Guidelines Program
Jerry Smrchek, Ph.D.

Food and Drug Administration
Office of Science
★ Suzanne Fitzpatrick, Ph.D., D.A.B.T.
Center for Drug Evaluation and Research
◊ Abigail C. Jacobs, Ph.D.
Paul C. Brown, Ph.D.
Center for Devices and Radiological Health
Melvin E. Stratmeyer, Ph.D.
Center for Biologics Evaluation and Research
Richard McFarland, Ph.D., M.D.
Ying Huang, Ph.D.
Center for Food Safety and Nutrition
David G. Hattan, Ph.D.
Robert L. Bronaugh, Ph.D.
Center for Veterinary Medicine
Devaraya Jagannath, Ph.D.
M. Cecilia Aguila, D.V.M.
National Center for Toxicological Research
William T. Allaben, Ph.D.
Paul Howard, Ph.D.
Office of Regulatory Affairs
Lawrence A. D'Hoostelaere, Ph.D.

National Cancer Institute
★ Alan Poland, M.D.
◊ T. Kevin Howcroft, Ph.D.

National Institute of Environmental Health Sciences
★ William S. Stokes, D.V.M., D.A.C.L.A.M.
◊ Raymond R. Tice, Ph.D.
Rajendra S. Chhabra, Ph.D., D.A.B.T
Jerrold J. Heindel, Ph.D.

National Institute for Occupational Safety and Health
★ Paul Nicolaysen, V.M.D.
◊ K. Murali Rao, MD, Ph.D.

National Institutes of Health
★ Margaret D. Snyder, Ph.D.

National Library of Medicine
◊ Jeanne Goshorn, M.S.

Occupational Safety and Health Administration
★ Surender Ahir, Ph.D

★ Principal Agency Representative
◊ Alternate Principal Agency Representative

Acknowledgements

The following individuals are acknowledged for their contributions to the ICCVAM in vitro pyrogen test method evaluation process.

Pyrogenicity Working Group (PWG)

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)

U.S. Food and Drug Administration
Center for Biologics Evaluation and Research
Mustafa Akkoyunlu, M.D., Ph.D.
Pankaj Amin
Christine Anderson
Kimberly Benton, Ph.D.
Joseph George, Ph.D.
Christopher Joneckis, Ph.D.
Richard McFarland, Ph.D, M.D. (Chair)
Center for Drug Evaluation and Research
David Hussong, Ph.D.
Abigail Jacobs, Ph.D.
Robert Mello, Ph.D.
Amy Rosenberg, M.D.
Daniela Verthelyi, M.D., Ph.D.
Jiaqin Yao, Ph.D.
Center for Devices and Radiological Health
Raju Kammula, D.V.M., Ph.D., D.A.B.T.
Ramesh Panguluri
Center for Food Safety and Nutrition
Penelope Rice, Ph.D.

U.S. Environmental Protection Agency
Office of Science Coordination and Policy
Karen Hamernik, Ph.D.

National Institute of Environmental Health Sciences
William Stokes, D.V.M., D.A.C.L.A.M.
Raymond Tice, Ph.D.

U.S. Department of Agriculture
Jodie Kulpa-Eddy, D.V.M.

Liaison, European Centre for the Validation of Alternative Methods
Marlies Halder, Ph.D.[1]

[1]Dr. Halder was responsible for the trial plan of the catch-up validation study, but she was not involved in the evaluation of the data.

National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

National Institute of Environmental Health Sciences

William Stokes, D.V.M., D.A.C.L.A.M.	Director; Project Officer
Raymond Tice, Ph.D.	Deputy Director
Deborah McCarley	Special Assistant; Asst. Project Officer

NICEATM Support Contract Staff (Integrated Laboratory Systems [ILS], Inc.)

David Allen, Ph.D.	Principal Investigator
Douglas Winters, M.S.	Project Manager
Neepa Choksi, Ph.D.	Senior Staff Toxicologist
Judy Strickland, Ph.D., D.A.B.T.	Senior Staff Toxicologist
Frank Deal, M.S.	Staff Toxicologist
Elizabeth Lipscomb, Ph.D.	Staff Toxicologist
Thomas Burns, M.S.	Senior Project Coordinator/Technical Writer
Michael Paris	Senior Project Coordinator/Technical Writer
Patricia Ceger, M.S.	Project Coordinator/Technical Writer
James Truax, M.A.	Project Coordinator/Technical Writer
Catherine Sprankle	Communications Specialist/Web Developer
Linda Litchfield	Meeting Planner and Coordinator

Consultants for ILS, Inc.

Joe Haseman, Ph.D.	Statistical Consultant
Errol Zeiger, Ph.D., J.D.	Editorial Consultant

Preface

Endotoxin, a bacterial pyrogen also known as lipopolysaccharide, is an integral component of the Gram-negative bacterial cell membrane. Endotoxin directly interacts with host monocytoid cells to induce the release of a variety of proinflammatory cytokines (e.g., interleukin [IL]-1β, IL-6, tumor necrosis factor-α). In addition to an initial febrile reaction, excessive release of these cytokines during Gram-negative bacterial sepsis can lead to multiple organ failure and death. For this reason, it is critical that parenteral pharmaceuticals, fluids for injection, medical devices, and human biological products be properly and accurately evaluated for the presence of endotoxin prior to their clinical or veterinary use. The original pyrogen test, the rabbit pyrogen test (RPT), was developed in 1941 to limit to an acceptable level the risks of febrile reaction in the patient following administration of, or contact with, the product of concern. While the RPT continues to serve this purpose well, an endotoxin test using a hemolymph extract (i.e., "blood") from the horseshoe crab (i.e., the bacterial endotoxin test [BET]) was developed in the early 1970's as an *in vitro* alternative to the RPT for the detection of Gram-negative endotoxin. In 1980, the United States (U.S.) Food and Drug Administration (FDA) published guidelines for use of the BET as an end-product test for human and animal drug products. The U.S., European, and Japanese Pharmacopeias currently recognize both test methods for pyrogen testing (i.e., RPT and BET). The BET is recognized for its sensitivity to the presence of endotoxins from Gram-negative bacteria, but it also has some limitations, including its inability to respond to non-endotoxin pyrogens, as well as its susceptibility to interference from certain types of materials (e.g., products with high protein and lipid levels, glucans). In contrast, the RPT is capable of detecting both endotoxin and non-endotoxin pyrogens.

More recent efforts have focused on the development of *in vitro* test systems that might achieve or exceed the sensitivity of the BET and the RPT. Test systems based on the activation of human monocytes *in vitro* have been developed that take advantage of the role of these cells in the fever response. The European Centre for the Validation of Alternative Methods (ECVAM), a unit of the Institute for Health and Consumer Protection at the European Commission's Joint Research Centre, conducted a validation study to independently evaluate the usefulness of six *in vitro* pyrogen test methods. The study was financed by the European Commission within the 5th Framework Programme of Directorate General Research and was recently published (Hoffmann et al. 2005a). Since two tests based on the acute monocyte leukemia cell line THP-1 did not meet the validation criteria, they are not included in the peer review. In 2004, the University of Konstanz (Germany) carried out catch-up validation studies of two tests using Cryopreserved whole blood (Cryo WB/IL-1β) or blood cells (cryopreserved or fresh peripheral blood mononuclear cells [PBMC]/IL-6), the results of which were recently published (Schindler et al. 2006).

Based on these studies, in June 2005, ECVAM submitted background review documents (BRDs) for five of these test methods, which were proposed as replacements for the RPT, to the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM). The five test methods are:

- The Human Whole Blood (WB)/IL-1β *In Vitro* Pyrogen Test

- The Human WB/IL-1β *In Vitro* Pyrogen Test: Application of Cryo Human WB
- The Human WB/IL-6 *In Vitro* Pyrogen Test
- The Human PBMC/IL-6 *In Vitro* Pyrogen Test
- The Monocytoid Cell Line Mono Mac 6/IL-6 *In Vitro* Pyrogen Test

For simplicity, the submitted studies are referred to collectively as the ECVAM validation study in this document.

ICCVAM, which is charged with coordinating the technical evaluations of new, revised, and alternative test methods with regulatory applicability (ICCVAM Authorization Act of 2000, [42 U.S. Code 285*l*-3, available at http://iccvam.niehs.nih.gov/docs/about_docs/PL106545.htm]), unanimously agreed that the five submitted *in vitro* test methods should have a high priority for evaluation. An ICCVAM Pyrogenicity Working Group (PWG) was established to work with the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) to carry out these evaluations. The PWG consists of knowledgeable scientists from ICCVAM member agencies. The PWG functions included reviewing draft test method BRDs, recommending proposed performance standards, identifying and recommending scientists for independent peer review panels, preparing questions for expert or peer review Panels, developing ICCVAM draft test method recommendations regarding the usefulness and applicability of the alternative test methods for regulatory testing, and recommending necessary future validation studies. ICCVAM and NICEATM also collaborate closely with ECVAM. Accordingly, an ECVAM liaison was designated for the ICCVAM PWG to provide additional clarification and information during the evaluation and review process.

NICEATM, which administers the ICCVAM and provides scientific support for ICCVAM activities, subsequently prepared a comprehensive draft BRD containing all of the information and data from the validation studies for each of the five *in vitro* test methods. A request for any other data and information on these test methods was made through a 2005 *Federal Register* (*FR*) request (Vol. 70, No. 241, pp. 74833-74834, December 16, 2005; available at http://ntp-apps.niehs.nih.gov/iccvampb/searchFR.cfm), through the ICCVAM electronic mailing list, and through direct requests to over 100 interested stakeholders. No additional data or information was submitted in response to this request.

The draft BRD was made publicly available on the NICEATM-ICCVAM website (http://iccvam.niehs.nih.gov). Comments from the public and scientific community were welcomed and were provided to the Panel and made available on the NICEATM-ICCVAM website (see *FR* notice [Vol. 71, No. 238, pp. 74533-74534, December 12, 2006], available at http://iccvam.niehs.nih.gov).

The independent review of the usefulness and limitations of the five test methods took place in a public meeting of the independent peer review panel (Panel) on February 6, 2007 at the National Institutes of Health in Bethesda, Maryland. The Panel considered the information and data available in the draft BRD. The Panel's independent peer review report was then made available for public comment on the NICEATM-ICCVAM website (see *FR* notice [Vol. 72, No. 89, pp. 26395-26396, May 9, 2007], available at http://iccvam.niehs.nih.gov). Following the Panel meeting, ICCVAM and the PWG considered the Panel's report and

public comments, and prepared this final BRD. ICCVAM and the PWG also considered the Panel's report, comments from the public and from the Scientific Advisory Committee on Alternative Toxicological Methods, and information in this BRD, and prepared final test method recommendations that will be provided to U.S. Federal agencies and made available to the public. These final recommendations are included in the ICCVAM Test Method Evaluation Report, which is available at http://iccvam.niehs.nih.gov/methods/pyrogen/pyrogen.htm, in accordance with the ICCVAM Authorization Act of 2000.

We acknowledge the ECVAM scientists who participated in the management of the validation studies and who prepared the ECVAM BRDs. We especially acknowledge Dr. Marlies Halder, ECVAM Liason to the PWG, for valuable information and comments throughout the review process. The efforts of many individuals who contributed to the preparation of the ICCVAM BRD are also gratefully acknowledged. These include Drs. David Allen and Elizabeth Lipscomb, Bradley Blackard, Catherine Sprankle, James Truax, and Doug Winters of Integrated Laboratory Systems, Inc., the NICEATM support contractor, as well as the members of the ICCVAM PWG and ICCVAM representatives who subsequently reviewed and provided comments throughout the process leading to this final version. We also want to thank Dr. Raymond Tice, Deputy Director of NICEATM, for his coordination efforts for this project. Finally, we want to recognize the excellent leadership of the PWG Chair, Dr. Richard McFarland, FDA.

William S. Stokes, D.V.M., D.A.C.L.A.M.
Rear Admiral, U.S. Public Health Service
Director, NICEATM
Executive Director, ICCVAM

Marilyn Wind, Ph.D.
U.S. Consumer Product Safety Commission
Chairman, ICCVAM

[This Page Intentionally Left Blank]

EXECUTIVE SUMMARY

This Background Review Document (BRD), prepared by the Interagency Coordinating Committee for the Validation of Alternative Methods (ICCVAM), provides a comprehensive description and analyses of the data and information supporting the validity of five *in vitro* pyrogen test methods. This BRD includes data from previously conducted validation studies and from previously published and unpublished data. The test methods are:

- The Human Whole Blood (WB)/Interleukin (IL)-1β *In Vitro* Pyrogen Test
- The Human WB/IL-1β *In Vitro* Pyrogen Test: Application of Cryopreserved (Cryo) Human WB
- The Human WB/IL-6 *In Vitro* Pyrogen Test
- The Human Peripheral Blood Mononuclear Cell (PBMC)/IL-6 *In Vitro* Pyrogen Test
- The Monocytoid Cell Line Mono Mac 6 (MM6)/IL-6 *In Vitro* Pyrogen Test

The validation studies evaluated the test methods for their ability to detect the presence of Gram-negative endotoxin that had been spiked into a range of injectable pharmaceuticals. This ICCVAM BRD provides information and data that support the current validation status of the *in vitro* pyrogen test methods. It discusses what is known about their relevance[1] and reliability[2], the types of substances tested, and the standardized test method protocols used to generate data for each test method.

Information in this ICCVAM BRD is based on data from five individual BRDs submitted by the European Centre for the Validation of Alternative Methods (ECVAM), a unit of the Institute for Health and Consumer Protection at the European Commission's Joint Research Centre (see **Appendix A**), to the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods. The ECVAM BRDs were prepared according to the ICCVAM submission guidelines (ICCVAM 2003). The ECVAM BRDs will also help agencies to assess whether the proposed test methods are acceptable for regulatory applications. Each ECVAM BRD summarizes the validation studies conducted for an individual *in vitro* pyrogen test method. This ICCVAM BRD compares and contrasts the performance of these five test methods.

This ICCVAM BRD also summarizes information from published studies and additional unpublished data provided by ECVAM. **Section 9.0** of this document discusses *in vitro* pyrogen test method studies that could not be included in the performance analyses because appropriate study details, test method results, or *in vivo* rabbit pyrogen test (RPT) reference data were not available. An online literature search for additional data on the proposed *in vitro* pyrogen test methods identified nineteen studies that contained relevant data. ECVAM also provided additional unpublished data in response to a request for additional information related to the validation studies (**Appendices B** and **C**).

---

[1]Relevance is the extent to which a test method correctly predicts or measures an effect, and includes the "accuracy" or "concordance" of the method.

[2]Reliability is a measure of how well a test method can be reproduced at different times and in different laboratories. It is assessed by calculating reproducibility both within and among laboratories and repeatability within laboratories.

An independent peer review panel (Panel) assessed the ICCVAM BRD for completeness and any errors or omissions. The Panel also evaluated the validation status of the proposed test methods in the ICCVAM BRD.

The *in vitro* pyrogen test methods discussed in this BRD measure release of the pro-inflammatory cytokines IL-1β or IL-6 in response to exposure to Gram-negative endotoxin. The test methods use monocytoid cells contained in WB, isolated PBMCs, or the MM6 cell line. No data were provided from the validation studies supporting the usefulness of these test methods for pyrogens other than endotoxins.

ICCVAM surveyed regulatory agencies in the United States (U.S.) to determine whether any of the proposed *in vitro* test methods have been considered for regulatory use where submission of test data is required. Regulatory practice in the U.S. and in the European Union is to accept pyrogen test method data for a specific product after the test method has been validated for that specific product. The ECVAM BRDs note that the U.S. Food and Drug Administration has accepted data from the PBMC test developed by Novartis and Baxter Healthcare. In this instance, the PBMC test results were used in conjunction with RPT and Bacterial Endotoxin Test data to support the safety testing of a single specific drug product (New Drug Application Number 16-267/S-037).

The predominant difference between the *in vitro* pyrogen test methods is the type of cells used. The following basic steps are consistent among all methods:

- Interference testing is performed to verify that a test substance does not interfere with either the cell system used or with the specific cytokine-specific enzyme-linked immunosorbent assay (ELISA).
- The test substance is mixed with a suspension of human-derived blood cells.
- The concentration of of the specific proinflammatory cytokine (e.g., IL-1β, IL-6) is measured using an ELISA, and is compared to the response curve of an endotoxin standard.
- An internationally accepted endotoxin standard (World Health Organization-lipopolysaccharide [WHO-LPS] 94/580 *Escherichia coli* [*E. coli*] O113:H10:K-), or an endotoxin standard that has been calibrated against this standard, is used to generate the standard response curve for the assay. The endotoxin activity of a test substance is calculated by comparing the induced cytokine release with that induced by the endotoxin standard.
- A product "passes" (i.e., is considered negative for endotoxin pyrogen activity) if the cytokine response to the test substance is less than that induced by 0.5 endotoxin units/mL (EU/mL).

The ability of the *in vitro* pyrogen test methods to correctly identify the presence of Gram-negative endotoxin was evaluated using 10 parenteral pharmaceuticals spiked with endotoxin (WHO-LPS 94/580 *E. coli* O113:H10:K-). Each drug, spiked with four concentrations of endotoxin, was tested once in three different laboratories. As indicated in **Table 1**, analysis of the five *in vitro* test methods indicated that accuracy among the test methods ranged from 81% to 93%, sensitivity ranged from 89% to 99%, specificity ranged

from 89% to 99%, specificity ranged from 81% to 97%, false negative rates[3] ranged from 1% to 27%, and false positive rates[4] ranged from 3% to 23%.

Table 1 Accuracy of *In Vitro* Pyrogen Test Methods[1]

Test Method	Accuracy[2]	Sensitivity[3]	Specificity[4]	False Negative Rate[5]	False Positive Rate[6]
Cryo WB/IL-1β	92% (110/120)	97% (75/77)	81% (35/43)	3% (2/77)	19% (8/43)
MM6/IL-6	93% (138/148)	96% (85/89)	90% (53/59)	5% (4/89)	10% (6/59)
PBMC/IL-6	93% (140/150)	92% (83/90)	95% (57/60)	8% (7/90)	5% (3/60)
PBMC/IL-6 (Cryo)[7]	87% (130/150)	93% (84/90)	77% (46/60)	7% (6/90)	23% (14/60)
WB/IL-6	92% (136/148)	89% (79/89)	97% (57/59)	11% (10/89)	3% (2/59)
WB/IL-1β (Tube)	81% (119/147)	73% (64/88)	93% (55/59)	27% (24/88)	7% (4/59)
WB/IL-1β (96-well plate)[8]	93% (129/139)	99% (83/84)	84% (46/55)	1% (1/84)	16% (9/55)

Abbreviations: Cryo = Cryopreserved; EU/mL = Endotoxin units per milliliter; IL = Interleukin; MM6 = Mono Mac 6; PBMC = Peripheral blood mononuclear cells; WB = Whole blood

[1]Data shown as a percentage (number of correct runs/total number of runs), based on results of 10 parenteral drugs tested in each of three different laboratories. Samples of each drug were tested with or without being spiked with a Gram-negative endotoxin standard (0, 0.25, 0.5, or 1.0 EU/mL, with 0.5 EU/mL tested in duplicate).

[2]Accuracy = the proportion of correct outcomes (positive and negative) of a test method.

[3]Sensitivity = the proportion of all positive substances that are classified as positive.

[4]Specificity = the proportion of all negative substances that are classified as negative.

[5]False negative rate = the proportion of all positive substances that are falsely identified as negative.

[6]False positive rate = the proportion of all negative substances that are falsely identified as positive.

[7]A modification of the PBMC/IL-6 test method that uses Cryo PBMCs.

[8]A modification of the WB/IL-1β test method that uses 96-well plates instead of tubes for the test substance incubation.

The RPT and *in vitro* pyrogen test results can be compared if the same substance is tested using both the *in vivo* RPT and *in vitro* methods (i.e., parallel testing data). However, because no RPT data were generated with the same test samples used in the *in vitro* test methods, the accuracy of the *in vitro* test results could not be compared directly with that of the RPT.

The limitations of these five *in vitro* test methods have not been fully evaluated. For this reason, product-specific validation will be necessary to establish if a particular test substance or material is appropriate for evaluation using these *in vitro* test methods. One identified limitation of the *in vitro* test methods is the lack of data to determine their responses to, and suitability for, pyrogens other than endotoxins that are currently detected by the RPT. However, a potential advantage of these *in vitro* test methods is that they are derived from

[3]False negative rates reflect a failure of the *in vitro* test method(s) to identify Gram-negative endotoxin spiked into a test substance at the threshold concentration (0.5 EU/mL) established based on historical data from the RPT.

[4]False positive rates reflect that the *in vitro* test method(s) identified the presence of Gram-negative endotoxin when it was not present.

human tissues, which avoids potential uncertainty associated with cross-species extrapolation.

Repeatability within individual laboratories was determined for each *in vitro* test method, using saline and various endotoxin spikes (0.06 to 0.5 EU/mL) to evaluate the closeness of agreement among optical density (OD) readings for cytokine measurements at each concentration. Up to 20 replicates per concentration were tested, and results indicated that variability in OD measurements increased with increasing endotoxin concentration. However, the variability was low enough that the threshold for pyrogenicity could still be detected (i.e., the 0.5 EU/mL spike concentration could still be distinguished from the lower concentrations).

Reproducibility within individual laboratories was evaluated using three marketed pharmaceuticals spiked with various concentrations of endotoxin. Three identical, independent runs were conducted in each of the three testing laboratories, with the exception of the Cryo WB/IL-1β test method[5]. The correlations (expressed as percentage of agreement) between pairs of the independent runs (i.e., run 1 vs. run 2; run 1 vs. run 3; run 2 vs. run 3) were determined, and the mean of these three values was calculated. Agreement between two runs within a single laboratory ranged from 75% to 100%, with mean values ranging from 83% to 100%. Agreement across three runs within a single laboratory ranged from 75% to 100%.

Reproducibility across all laboratories was evaluated in two different studies in which each run from one laboratory was compared to all other runs of another laboratory. The proportion of equally qualified samples provided a measure of reproducibility. In the first reproducibility study, three marketed pharmaceutical products were spiked with either saline control or various concentrations of endotoxin, and each sample was tested in triplicate in each of three different laboratories, except for Cryo WB/IL-1β. In the catch-up validation study of Cryo WB/IL-1β, each sample was tested once in each laboratory. The agreement across the three laboratories for each test method ranged from 58% to 86%[6], depending on the test method used, and 92% for the Cryo WB/IL-1β test method.

In the second study, reproducibility was determined using the results from the 10 drugs used in the accuracy analysis. Each drug was spiked with four concentrations of endotoxin and tested once in each of three laboratories. The agreement across three laboratories for each test method ranged from 57% to 88%, depending on the test method used. The extent and order of agreement among laboratories were the same for both studies: the WB/IL-1β test method showed the least agreement (57% to 58%), and the Cryo WB/IL-1β test method showed the most (88% to 92%).

This ICCVAM BRD provides a comprehensive summary of available data used to determine the usefulness and limitations of five *in vitro* pyrogen test methods for detecting Gram-negative endotoxin. It discusses what is currently known about their relevance and reliability,

---

[5]The ECVAM Cryo WB/IL-1 test method BRD states that there was no direct assessment of intralaboratory reproducibility because such an evaluation was performed in the WB IL-1 (fresh blood) test method, and the authors assumed that variability is not affected by the change to cryopreserved blood.

[6]However, a modification of the WB/IL-1 test method (using 96-well plates for the test substance incubation) resulted in agreement among laboratories of 83% to 92% when tested once in each laboratory.

the types of the substances tested, and the standardized test method protocols used to generate data for each test method. The information in this BRD was used by ICCVAM to finalize its recommendations for test method uses, standardized test method protocols, and future studies to further characterize the usefulness and limitations of these test methods. These test method recommendations will be provided to U.S. Federal agencies for consideration, in accordance with the ICCVAM Authorization Act of 2000 (42 U.S. Code § 285*l*-2 through 285*l*-5), available at: http://iccvam.niehs.nih.gov/about/about_ICCVAM.htm. Agency responses to ICCVAM will be available on the NICEATM-ICCVAM website (http://iccvam.niehs.nih.gov) 180 days after agency receipt of the recommendations.

[This Page Intentionally Left Blank]

1.0 Introduction And Rationale For The Proposed Use Of *In Vitro* Pyrogen Test Methods

1.1 Introduction

1.1.1 Historical Background of In Vitro Pyrogen Test Methods and the Rationale for Their Development

A brief summary of the historical development of the five *in vitro* pyrogen test methods was provided in Section 1.1.1 of each Background Review Document (BRD) provided by the European Centre for the Validation of Alternative Methods (ECVAM), a unit of the Institute for Health and Consumer Protection at the European Commission's Joint Research Centre. These BRDs were provided to the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and are included in **Appendix A**[1]. This section includes supplementary information and provides a context for United States (U.S.) regulatory considerations.

Pyrogenic substances (i.e., substances that induce fever) may originate from a variety of biological or synthetic/manufacturing sources. They may also be released from microbiological organisms such as bacteria, viruses, and fungi during cell death or following immunological attack (i.e., cell damage or death due to a local or systemic immune response). One of the most potent pyrogenic materials is bacterial endotoxin, which is an outer membrane component of the Gram-negative bacteria cell wall. Pyrogens may also be found in processing and packaging materials, chemicals, raw materials, or equipment used during the manufacturing of parenteral drugs or medical devices. The presence of endotoxins in otherwise sterile biological preparations such as parenteral drugs suggests the presence of past or current bacterial contamination.

The induction of fever by these pyrogenic substances is a complex process and multiple mechanisms are thought to be involved. It is likely that the specific pathway, or combinations of pathways, involved in the production of a fever response depends on a number of variables (e.g., the properties of the pyrogenic substance and the route of administration). In general, pyrogenic substances cause leukocytes (i.e., neutrophils, monocytes/macrophages, and lymphocytes) to release cytokines (e.g., interleukin [IL]-1β, IL-6, and Tumor Necrosis Factor-α [TNF-α]) that act as endogenous pro-inflammatory mediators, often referred to as "endogenous pyrogens" (Dinarello 1999). Once released, these cytokines act on the central nervous system to promote the synthesis of prostaglandins, ultimately producing of a fever response (Dinarello 1999; Netea et al. 2000). These cytokines have been shown to be associated with the fever response induced by pyrogenic substances in both humans and rabbits (Dinarello 1999). Certain bacterial products (e.g., endotoxin) can also stimulate cytokine production directly through the activation of Toll-like receptors (TLRs) (Dinarello 1999; Netea et al. 2000).

[1]References to ECVAM BRD sections are in normal type to distinguish them from references to ICCVAM BRD sections, which are in **boldface** type.

The translation of released cytokines into a fever response is largely mediated by circumventricular organs. These small neuronal cell groups allow neurons to come in contact with a variety of circulating substances directly from the bloodstream, which are thought to control the febrile response through projections to sites in the hypothalamus and brain stem (Saper and Breder 1994; Dinarello 1999; Beutler and Rietschel 2003).

The U.S., European, and Japanese Pharmacopeias currently recognize two test methods for pyrogen testing, the *in vivo* rabbit pyrogen test (RPT) and the *in vitro* bacterial endotoxin test (BET), also referred to as the *Limulus* amebocyte lysate (LAL) test. The BET is accepted because of its sensitivity to the presence of Gram-negative endotoxins. However, the test method has well documented limitations, including its inability to respond to non-endotoxin pyrogens, as well as its susceptibility to interference from certain types of materials (e.g., high protein and lipid levels, glucans). In contrast, the RPT is capable of detecting both endotoxin and non-endotoxin pyrogens. However, disadvantages of the RPT include the need for interspecies extrapolation from rabbits to humans.

In 2002, a total of 243,838 rabbits were used in the U.S. for all research and testing purposes, of which 6,324 rabbits were reported as experiencing more than slight or momentary pain and/or distress where anesthetics, analgesics, or tranquilizers could not be administered for scientific reasons (U.S. Department of Agriculture [USDA] 2002). Eight of these cases were specifically attributed to pyrogenicity testing, presumably based on induction of a fever response (USDA 2002). Thus, although the potential for more than slight or momentary pain and/or distress exists for pyrogenicity testing when a fever response is induced, it does not appear that a fever response is common. In the European Union (EU), approximately 313,000 total rabbits were used for all scientific purposes in 2005 (CEC 2007). Of these, approximately 276,000 rabbits were used for pharmaceutical products and medical device testing (i.e., either research and development, production and quality control, or toxicological and other safety evaluations). Although the number of rabbits specifically used for pyrogenicity testing was not reported, it is likely that this number is significantly less than the total of 276,000. Additional animal use numbers, including data reported from Canada and the United Kingdom (U.K.) are summarized in **Section 10.1**.

An *in vitro* test system that combines the sensitivity of the BET with the wide range of pyrogens detectable by the RPT would be an obvious improvement for pyrogen testing. With this intention, test systems based on the activation of human monocytes *in vitro* were developed that take advantage of an increased understanding of the biological mechanisms responsible for the human fever reaction (Dinarello 1999). Initial efforts focused on peripheral blood mononuclear cells (PBMC), which release proinflammatory cytokines when exposed to endotoxin (Duff and Atkins 1982; Dinarello et al. 1984). A number of similar test systems, using either whole blood (WB), PBMCs, or monocytoid cell lines (e.g., Mono Mac 6 [MM6], THP-1) were subsequently developed (Tsuchiya et al. 1980; Poole et al. 1988; Ziegler-Heitbrock et al. 1988; Hartung and Wendel 1996; Hartung et al. 2001; Poole et al. 2003; Gaines Das et al. 2004). Five test systems developed from human monocytoid cells were selected by ECVAM for prevalidation and validation studies with the intent of comparing their effectiveness for replacing the RPT and thereby eliminating the use of rabbits for pyrogen testing. The results of these studies have been published (Hoffmann et al. 2005a; Schindler et al. 2006). The five tests selected were:

- The Human WB/IL-1β *In Vitro* Pyrogen Test[2]

- The Human WB/IL-1β *In Vitro* Pyrogen Test: Application of Cryopreserved (Cryo) Human WB
- The Human WB/IL-6 *In Vitro* Pyrogen Test
- The Human PBMC/IL-6 *In Vitro* Pyrogen Test[2]
- The MM6/IL-6 *In Vitro* Pyrogen Test

1.1.2 Peer Reviews of In Vitro Pyrogen Test Method Validation Studies

The ECVAM-sponsored validation studies of each of these *in vitro* test methods have been the subject of a recent formal peer review convened by the ECVAM Scientific Advisory Committee (ESAC). Two members of the ESAC served as co-chairpersons for the review Panel, which consisted of five additional U.S. and European reviewers. These reviewers assessed the ability of each test method to serve as a complete replacement for the RPT. Based on this review, the ESAC declared that, "these tests have been scientifically validated for the detection of pyrogenicity mediated by Gram-negative endotoxins, and quantification of this pyrogen, in materials currently evaluated and characterized by RPTs." Although the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) requested the ESAC peer review report, this document is not currently permitted by ECVAM to be publicly disseminated.

This BRD was prepared for an ICCVAM independent peer review panel (Panel) to evaluate these *in vitro* pyrogen test methods and to consider the ICCVAM draft recommendations for each *in vitro* test method. Because individual BRDs for each test method were provided by ECVAM, the ICCVAM BRD provides information that was common to all five *in vitro* test methods and references the appropriate sections of the ECVAM BRDs for specifics related to individual test methods. The recommendations of the ICCVAM Panel, combined with the information and analyses presented in the ICCVAM and ECVAM BRDs and any comments by the public or the Scientific Advisory Committee on Alternative Toxicological Methods were considered by ICCVAM prior to making its final recommendations on the usefulness and limitations of each test method, the proposed standardized test method protocols, performance standards, and any additional studies considered necessary to further develop or characterize any or all of these *in vitro* test methods.

1.2 Regulatory Rationale and Applicability

1.2.1 Current Regulatory Testing Requirements and ICCVAM Prioritization Criteria

This section reviews and summarizes the extent to which the five ICCVAM prioritization criteria (ICCVAM 2003) apply to the *in vitro* pyrogen test methods under consideration.

Criteria 1. The extents to which the proposed test methods are (a) applicable to regulatory testing needs and (b) applicable to multiple agencies/programs.

---

[2]As indicated in the ECVAM BRDs for the WB/IL-1 and PBMC/IL-6 test methods, catch-up validation studies were conducted to evaluate the performance of the WB/IL-1 test method when using 96-well plates, and the PBMC/IL-6 test method when using cryopreserved PBMCs. The plating procedure (WB/IL-1) and the cryopreservation procedure (PBMC/IL-6) are the only differences in the test method protocols (see **Appendix A**). These modifications were not submitted by ECVAM as separate test methods, and are therefore not being considered as separate test methods in this BRD. However, where relevant, comparative information is provided (e.g., see **Table 2-1** and **Sections 6.1** and **7.2**).

Pyrogenicity testing is primarily used by regulatory authorities for end-product release of human and animal parenteral drugs, biological products, and medical devices. The results from these assays are used to limit, to an acceptable level, the risks of febrile reaction in the patient exposed to the product of concern by injection and/or implantation. As detailed in **Table 1-1**, the U.S. Food and Drug Administration (FDA) is the principal U.S. regulatory agency that requires pyrogenicity testing, with different Centers within the FDA regulating the affected products. The Center for Biologics Evaluation and Research, the Center for Drug Evaluation and Research, the Center for Devices and Radiological Health, and the Center for Veterinary Medicine require that human injectable drugs (including biological products), animal injectable drugs, and medical devices be tested for the presence of pyrogenic substances. The current U.S. legislation requiring the use of pyrogenicity testing is set forth in the Federal Food, Drug, and Cosmetic Act (U.S. Code [U.S.C.], Title 21, Chapter 9). In addition, the U.S. Pharmacopeia (USP) maintains sterility requirements for pharmaceuticals that include pyrogenicity testing. **Table 1-1** also shows the statutory protocol requirements used by each FDA Center, along with the comparable enabling legislation and statutory protocol requirements of the EU member nations.

Additionally, the prediction model described in the ECVAM BRDs is based on a pyrogen threshold dose of 0.5 EU/mL. While this level of detection would be sufficient for many parenteral drugs and medical devices, the endotoxin limit set by the FDA for intrathecal drugs and devices that contact cerebrospinal fluid is 0.06 EU/ml. In response to an ICCVAM Pyrogenicity Working Group (PWG) request for more data to support the use of these test methods for discriminating an endotoxin threshold dose lower than 0.5 EU/ml, ECVAM provided supplemental data (see question #6 in **Appendix B**).

Table 1-1 Summary of U.S. and European Legislation and Statutory Protocol Requirements for Pyrogenicity Testing

Agency	Regulated Products	Legislation	Statutory Protocol Requirements	Non-Governmental Standards
United States				
FDA-CBER	Biological products	Federal Food, Drug, and Cosmetic Act (U.S.C. Title 21, Chapter 9)	21 CFR 610.13 (FDA 2005)	USP30 NF25<85> (USP 2007a) USP30 NF25<151> (USP 2007b) ISO 10993-11 (ISO 2006)
FDA-CDER	Human parenteral pharmaceuticals			
FDA-CDRH	Medical devices			
FDA-CVM	Veterinary pharmaceuticals			
Europe				
EDQM	Human/veterinary parenteral pharmaceuticals, biological products, medical devices	Council Regulation (EEC) 230/9/93 Council Directive 93/39/EEC Council Directive 93/40/EEC	EP5.0 2.6.8 (EP 2005a) EP5.0 2.6.14 (EP 2005b)	ISO 10993-11 (ISO 2006)
EMEA				
Regulatory Authorities for Individual EU Countries				

Abbreviations: CBER = Center for Biologics Evaluation and Research; CDER = Center for Drug Evaluation and Research; CDRH = Center for Devices and Radiological Health; CFR = U.S. Code of Federal Regulations; CVM = Center for Veterinary Medicine; EDQM = European Directorate for the Quality of Medicines; EEC = European Economic Community; EMEA = European Medicines Agency; EP = European Pharmacopeia; EU = European Union; FDA = U.S. Food and Drug Administration; ISO = International Standards Organization; NF = National Formulary; U.S.C. = United States Code; USP = U.S. Pharmacopeia

Criteria 2. Warranted, based on the extent of expected use or application and impact on human, animal, or ecological health.

The proposed test methods are intended to replace a method that is used extensively in pharmaceutical, biological product, and medical device development and registration (i.e., the RPT).

Criteria 3: The potential for the proposed test methods, compared to current test methods accepted by regulatory agencies, to (a) refine animal use (decrease or eliminate pain and distress), (b) reduce animal use, or (c) replace animal use.[3]

The two most common pyrogen tests presently used (i.e., RPT and BET) require the use of animals. The RPT is performed in rabbits that can be maintained and reused (under certain circumstances) for multiple tests. According to USP30 NF25<151> (USP 2007b), rabbits may not be reused more than once every 48 hours (hr) after a negative test, not less than two

[3]*Refinement alternative* is defined as a new or revised test method that refines procedures to lessen or eliminate pain or distress to animals, or enhances animal well-being. *Reduction alternative* is defined as a new or revised test method that reduces the number of animals required. *Replacement alternative* is defined as a new or revised test method that replaces animals with nonanimal systems or one animal species with a phylogenetically lower one (e.g., a mammal with an invertebrate) (ICCVAM 1997).

weeks following either (1) a maximum rise of 0.6°C or more, or (2) an animal is included in a test with a substance that is classified as pyrogenic. The BET is performed using hemolymph (the equivalent of blood, which requires drawing approximately 20% of the animal's total blood volume) obtained from *Limulus polyphemus* (horseshoe crabs). Although the donor horseshoe crabs are returned to the wild, some mortality (up to 15%) is associated with the procedure (Walls et al. 2002).

The need for horseshoe crab hemolymph has potentially been reduced with the development of recombinant Factor C, the endotoxin sensitive protein that initiates clotting in the traditional BET. This commercial product, which was originally cloned from the horseshoe crab (Ding et al. 1997; Ding and Ho 1998, 2001), is currently being compared to the BET for submission and inclusion in the USP.

It should also be noted that the FDA has accepted data from the PBMC test developed by Novartis and Baxter Healthcare, which in conjunction with RPT and BET results, were used to support the safety testing of a specific single drug product (New Drug Application Number 16-267/S-037).

Criteria 4: The potential for the proposed test methods to provide improved prediction of adverse health or environmental effects, compared to current test methods accepted by regulatory agencies.

Sufficient data are presented to allow an assessment of the performance of the proposed test methods relative to the RPT (see **Section 6.0**). Because these methods are conducted using cells of human origin, it is postulated that they may reflect the human physiological response better than the currently employed, non-human based methods (i.e., RPT and BET).

These *in vitro* test methods have also been found to be useful for detecting test substances that, in the absence of endotoxin contamination, have evoked an adverse response in patients (Marth and Kleinhappl 2002; Martis et al. 2005). For example, numerous cases of aseptic peritionitis in dialysis patients that were not febrile have been attributed to peptidoglycan contamination of the dialysate (Martis et al. 2005). This dialysate solution met all European and USP standards prior to product release, but the PBMC/IL-6 test method detected increased levels of IL-6 when the dialysate was tested following product recall (Martis et al. 2005). In Marth and Kleinhappl (2002), a case study of a vaccine that was approved for release by the Austrian health authorities, but later produced a fever response in humans, has been described. When this vaccine was subsequently tested in the WB/IL-1β test method, it produced a positive result (Marth and Kleinhappl 2002).

Criteria 5: The extent to which the test method provides other advantages (e.g., reduced cost and time to perform) compared to current methods.

As outlined in **Table 11-1**, cost estimates obtained from various contract laboratories that perform the RPT and from the only contract laboratory known to perform an *in vitro* pyrogen test indicate that the *in vitro* test methods are considerably more cost effective to perform than the RPT. With respect to time considerations, the *in vitro* test methods require two half-days (i.e., one before and one after the overnight incubation) to complete, provided that cryopreserved blood is available and that interference testing is not required. The RPT can be performed within one day. However, before using a rabbit for the first time in a RPT, it must

be conditioned by a sham test that includes all steps of pyrogenicity testing except for injection, according to USP30 NF25<151> (USP 2007b).

1.2.2 Intended Uses of the Proposed In Vitro Pyrogen Test Methods

The proposed test methods are intended as an end-product release test for the identification of Gram-negative endotoxin in human and animal parenteral drugs, biological products, and medical devices. The results from pyrogen testing are used to limit, to an acceptable level, the risks of febrile reaction to the injection and/or implantation of the product of concern.

1.2.3 Similarities and Differences in the Endpoints Measured by the Proposed Test Methods and the In Vivo Reference Test Method

The endpoint measured in the *in vitro* pyrogen test methods is release of proinflammatory cytokines, either IL-1β or IL-6, in response to a test substance challenge, depending on the specific cell type employed. As described in **Section 1.1.1**, the pathogenesis of fever is induced by bacterial products that stimulate the production of IL-1α, IL-1β, and TNF-α, which leads to secondary synthesis of IL-6 and subsequent induction of prostaglandin synthesis (Netea et al. 2000). Direct injection of either IL-1 or IL-6 in several species causes fever, but much higher concentrations of IL-6 are needed. For example, in the rabbit, up to 100-fold more IL-6 is needed to produce a fever compared to IL-1 (Dinarello 2004).

The RPT involves measuring the rise in body temperature evoked in rabbits by the intravenous (i.v.) injection of a test solution. Although there is no direct association between the endpoints measured in the *in vitro* test methods and the RPT, fever is mediated by proinflammatory cytokines and therefore, it is reasonable to postulate that the cytokine-inducing potential of a pyrogen should correlate with its pyrogenic potential (Nakagawa et al. 2002). Moreover, Nakagawa et al. (2002) evaluated the utility of *in vitro* pyrogen test methods for detecting and quantifying various pyrogens. For example, the authors demonstrated that the responsiveness of human WB cells to pyrogens was very similar to that of a subline of MM6 cells, where endotoxin treatment (1 ng/mL) resulted in the production of IL-6 (~1 ng/mL) and IL-1 (~0.1 ng/mL).

Because the RPT is based solely on a rise in body temperature, no data were found on proinflammatory cytokine levels in rabbits following injection with endotoxin to permit a direct comparison with the *in vitro* test methods.

1.2.4 Use of the Proposed Test Methods in an Overall Strategy of Hazard or Safety Assessment

As detailed in **Table 1-1**, current U.S. and European regulatory requirements exist to test pharmaceutical products, biological products, and medical devices for pyrogenicity. The pyrogen tests that are currently acceptable to regulatory authorities require the use of rabbits or horseshoe crab hemolymph. According to ECVAM, the *in vitro* test methods are intended to replace the RPT for the identification of pyrogens where: (a) the test material is incompatible with the BET or (b) the test material contains a non-endotoxin mediated pyrogen. However, as detailed in **Section 3.0**, only Gram-negative endotoxin was included in the validation study. Therefore, other types of pyrogens have not been adequately validated

(refer to **Section 1.1.2**)[4]. The extent to which the RPT is performed only for detecting the presence of endotoxin is not clear.

1.3 Scientific Basis for the *In Vitro* Pyrogen Test Methods

1.3.1 Purpose and Mechanistic Basis of the In Vitro Pyrogen Test Methods

The proposed methods are intended to detect pyrogens in parenteral pharmaceuticals, medical devices, and human biological products. These methods provide an *in vitro* model of the initiation of the human fever response by measuring proinflammatory cytokine release (i.e., IL-1β or IL-6) from human monocytes/monocytoid cells exposed to pyrogens. These proinflammatory cytokines are associated with the initiation of the *in vivo* fever response.

1.3.2 Similarities and Differences of Modes of Action Between the In Vitro Pyrogen Test Methods and the Fever Response in Humans and/or Rabbits

As detailed in **Section 1.2.3**, each of the five proposed *in vitro* test methods measure proinflammatory cytokine release (i.e., IL-1β or IL-6) from human monocytoid cells as an indicator of the presence of a pyrogenic substance. By comparison, the RPT measures a change in body temperature in rabbits over a specified time period following an i.v. injection of a test substance. Although the relative sensitivities of each species to Gram-negative endotoxins vary, the responses of humans, horseshoe crabs (via hemolymph gelatin), and rabbits to these pyrogens have been studied extensively, and test methods based on blood products or blood cells from each of these species appear to be capable of responding to pyrogens (Greisman and Hornick 1969; Cooper et al. 1971; Brunson and Watson 1974; Hoffman et al. 2005a). Several studies directly comparing the *in vitro* pyrogen test methods with either the RPT and/or BET are summarized in **Section 9.1**, Moseby et al. (2000), and in the ECVAM response to ICCVAM PWG questions (see question #1 in **Appendix B**).

The recent discovery and characterization of the TLR family, which recognizes a diverse range of molecules such as lipids, proteins, and nucleic acids derived from pathogens, has led to an enhanced understanding of the signaling pathways activated by endotoxin. More specifically, TLR-4 has been identified as the receptor directly utilized by endotoxin to elicit an immune response. Upon recognition of endotoxin, TLR-4 initiates a rapid and complex signaling cascade, which activates transcription factors (i.e., NF-κB, AP-1, and interferon regulatory factors) to produce proinflammatory cytokines and other immune modulators, thereby leading to a protective immune response (Ishii et al. 2005; Ishii and Akira 2006). It is important to recognize that this TLR-dependent production of proinflammatory cytokines is distinct from the endotoxin-induced synthesis of IL-1 that then converges on the same signaling pathway via the IL-1 receptor (Conti et al. 2004). In addition to endotoxin, TLR-4 recognizes numerous other microbial components such as respiratory syncytial virus proteins and anthrolysin O (Ishii et al. 2005, Ishii and Akira 2006). When proinflammatory cytokine mRNA levels (i.e., IL-1, IL-6, and TNF-α) were compared in response to various TLR-4 agonists, endotoxin induced the highest level of expression (Park et al. 2004).

[4]Additional information on testing of other types of pyrogens was also provided by ECVAM and is included in **Appendix B**.

It has been recognized for many years that humans are responsive to relatively low doses of endotoxin, whereas rodents require much higher doses to elicit a response. In recent years, these species differences have been attributed, in part, to structural differences in TLR-4. For this reason, caution should be used when extrapolating findings from other mammals to humans with respect to endotoxin and TLR-4 signaling (Stoll et al. 2006). Furthermore, TLR-4 mutations have been identified in mice and humans, and it is likely that such defects are associated with altered gene expression and increased susceptibility to infection (Norata et al. 2005; van Deventer 2000; von Aulock et al. 2003).

1.3.3 Range of Substances Amenable to the In Vitro Pyrogen Test Methods and Limits of These Methods

The proposed methods are intended for the identification of pyrogenic substances in parenteral pharmaceuticals, biological products, and medical devices. Because they are based on cultured human monocytes/monocytoid cells, they are considered capable of detecting both Gram-negative endotoxin and non-endotoxin-based pyrogens. While **Section 9.0**, Moesby et al. (2005), and the ECVAM response to ICCVAM PWG questions (see question #2 in **Appendix B**) provide a number of published studies demonstrating that the *in vitro* pyrogen test methods are able to detect non-endotoxin pyrogens, the ECVAM validation studies focused specifically on Gram-negative endotoxin due to the unavailability of standardized, non-endotoxin pyrogens (see **Section 3.0**).

Because these test methods measure the release of proinflammatory cytokines, drugs that are cytotoxic to blood cells or that induce a substantial proinflammatory response (e.g., IL-1 receptor antagonists, interferon [IFN]- γ , and rheumatic factors) are not amenable to testing by these methods (Hartung et al. 2001; Ishii et al. 2005; Ishii and Akira 2006). As described in **Section 2.0**, each test method includes an interference test to identify problematic test samples.

1.4 Validation of the *In Vitro* Pyrogen Test Methods

The ICCVAM Authorization Act of 2000 mandates that "[each] Federal Agency … shall ensure that any new or revised … test method … is determined to be valid for its proposed use prior to requiring, recommending, or encouraging [its use]." Validation is the process by which the reliability and relevance of an assay for a specific purpose are established (ICCVAM 1997). Relevance is defined as the extent to which an assay will correctly predict or measure the biological effect of interest (ICCVAM 1997). For the *in vitro* pyrogen test methods described in this ICCVAM BRD, relevance is restricted to how well the assays detect the presence of Gram-negative endotoxin. Reliability is defined as the reproducibility of a test method within and among laboratories and should be based on performance with a diverse set of substances that are representative of the types of chemical and product classes that are to be tested and the range of responses that needs to be identified. The validation process is designed to provide data and information that will allow ICCVAM to make recommendations on the applicability of a test method and U.S. Federal agencies to consider those recommendations in light of their regulatory mandates.

The first stage in the evaluation of a new test procedure is the preparation of a BRD that presents and evaluates the relevant data and information about the test method, including its

mechanistic basis, proposed uses, reliability, and performance characteristics (ICCVAM 1997). This ICCVAM BRD summarizes the available information on each of the five *in vitro* pyrogen test methods listed in **Section 1.1.1**.

Where adequate data are available, the qualitative and quantitative performance of the proposed alternative test method is evaluated, and its reliability is compared with the reliability of the currently accepted test method. This ICCVAM BRD will aid in identifying essential test method components that should be considered during the identification of a standardized protocol for use of the test method.

1.5 Search Strategies and Selection of Citations for the ICCVAM *In Vitro* Pyrogen Test Methods BRD

NICEATM conducted an online literature search for relevant information on the five *in vitro* pyrogen test methods using multiple internet databases (i.e., PubMed, SCOPUS, TOXLINE, Web of Science). Specifically, records were sought using various combinations of the terms: *in vitro*, WB, WB cells, PBMC, Mono Mac 6, MM6, endotoxin, lipopolysaccharide (LPS), pyrogen, LAL, BET, IL-1, and IL-6. This search was conducted to supplement and update the list of peer-reviewed publications related to *in vitro* pyrogen testing that was provided in the ECVAM BRDs. U.S., EU, and Japanese pyrogenicity test guidelines were obtained from relevant regulatory agencies via the internet or through direct requests. The resulting database of 370 references confirmed that the lists of references included in the ECVAM BRDs were complete and up-to-date.

2.0 *In Vitro* Pyrogen Test Method Protocol Components

2.1 Overview of How the *In Vitro* Pyrogen Test Methods Are Conducted

Although there are differences among the *in vitro* pyrogen test methods considered in this ICCVAM BRD, the basic procedural steps are consistent across all five methods:

- Interference testing is performed to verify that a test substance does not interfere with either the cell system used or with the specific cytokine-specific enzyme-linked immunosorbent assay (ELISA).
- The test substance is mixed with a suspension of human-derived blood cells.
- The concentration of the specific proinflammatory cytokine (e.g., IL-1β, IL-6) is measured using an ELISA, and is compared to the response curve of an endotoxin standard.
- An internationally accepted endotoxin standard (World Health Organization-LPS [WHO-LPS] 94/580 *Escherichia coli* [*E. coli*] O113:H10:K-), or an endotoxin standard that has been calibrated against this standard, is used to generate the standard response curve for the assay. The endotoxin activity of a test substance is calculated by comparing the induced cytokine release with that induced by the endotoxin standard.
- A product "passes" (i.e., is considered negative for endotoxin pyrogen activity) if the cytokine response to the test substance is less than that induced by 0.5 endotoxin units/mL (EU/mL).

2.2 Description and Rationale for the Test Method Components for Proposed Standardized Protocols

The standard operating procedures for each test method assessed in the ECVAM validation studies are provided as Appendix A of each ECVAM BRD. As indicated in **Section 2.1**, there are essential principles of each protocol that are common among the five methods reviewed. These include:

- Isolating and/or culturing human monocytoid cells (either included in WB, separated as a fraction [i.e., PBMCs], or as cell line [i.e., MM6])
- Performing interference testing with each substance
- Treating the cells in suspension with a test substance
- Collecting cytokine release data
- Evaluating the data in relation to the proposed prediction model

Table 2-1 provides a comprehensive comparison of the similarities and differences among the protocols for the five test methods. No rationale was provided for the use of WB in the various test methods; however, Poole et al. (2003) summarized several studies, which indicated that the monocytes present in diluted WB respond to pyrogen/endotoxin by releasing pyrogenic cytokines. The use of the MM6 cell line was justified based on mechanistic considerations and its response to endotoxins.

[This Page Intentionally Left Blank]

Table 2-1 *In Vitro* Pyrogen Test Method Components

	Test Method Component	WB/IL-1β[1]	Cryo WB/IL-1β	WB/IL-6	PBMC/IL-6[1]	MM6/IL-6
	Source of cells	Human WB	Human Cryo WB	Human WB	Human WB	MM6 cell bank (original cell line maintained by Prof. H. Ziegler-Heitbrock, U. Munich)
	Laboratory equipment	• CO_2 cell culture incubator (37°C, 5% CO_2, humidified) • Centrifuge • Consumables as specified in SOP (e.g., heparinized blood tubes for WB methods, 96-well plates or culture tubes, centrifuge tubes, microfuge tubes, pyrogen-free plastic-ware where specified, serological pipets, pyrogen-free hypodermic needles) • Data analysis software • Hemocytometer (e.g., PBMC and MM6 assays) • Laminar Flow Hood (Class II) • Liquid nitrogen, CO_2 freezers, or programmable freezers for cryopreservation methods (Cryo WB/IL-1β or Cryo PBMC) • Microscope, inverted (optional except for PBMC and MM6 assays) • Microtiter Plate Reader (450 nm with 600-690 nm reference filter for IL-1β or 500-590 nm reference filter for IL-6 measurements) • pH meter • Pipettors (8 to 12 multi-channels; 2 to 2000 µL adjustables; pyrogen-free tips (except for ELISA) • Vortex mixer • Water bath				
	Culture medium	None - WB is diluted with PFS	RPMI Complete Medium • RPMI 1640 (part of the Endosafe Kit for cryoblood) • no specific additives needed	None - WB is diluted with PFS	RPMI Complete Medium • RPMI 1640 • HSA • L-Glutamine (2 mM) • Penicillin/streptomycin	RPMI Medium[2] • RPMI 1640 medium • Bovine insulin (0.23 IU/mL) • HEPES (20 mM) • HIFCS (10% or 2%) • L-glutamine (2mM) • MEM non-essential amino acid solution (0.1 mM) • Oxaloacetic acid (1 mM) • Sodium pyruvate (1 mM)
	Other reagents	• Endotoxin standard • PFS • PFW • Validated IL-1β ELISA kit	• DMSO • Endotoxin standard • PFS • PFW • Validated IL-1β ELISA kit	• Endotoxin standard • PFS • PFW • Validated IL-6 ELISA kit	• Endotoxin standard • PFS • PFW • Trypan blue • Validated IL-6 ELISA kit	• DMSO • Endotoxin standard • PFS • PFW • Trypan blue • Validated IL-6 ELISA kit
Dose selection procedures		Interference testing performed to determine the lowest dilution of the test product necessary to achieve an acceptable endotoxin spike recovery (i.e., 50% to 200% recovery)[3]				
Endpoints measured		IL-1β release via ELISA		IL-6 release via ELISA		

	Test Method Component	WB/IL-1β[1]	Cryo WB/IL-1β	WB/IL-6	PBMC/IL-6[1]	MM6/IL-6
Exposure of the test substance	**Pre-test preparation of cells**	Collect WB, heparinize, and use within 4 hr Plate Method: same collection procedure	• Collect WB, heparinize, and cryopreserve according to the Konstanz or PEI method • Prior to testing, thaw WB at 37°C for 15 min	Collect WB, heparinize, and use within 4 hr	• Collect WB and isolate PBMCs by centrifugation • Resuspend PBMCs in RPMI-C ($1x10^6$ cells/mL) (use PBMCs within 4 hr of initial WB collection)	• Incubate MM6 cells ($4x10^5$ cells/mL media) for 24 hr • Resuspend cells ($2.5x10^6$ cells/mL)[3] prior to testing
	Application of the test substance	Tube method: In a microfuge tube mix 1000µL PFS+100µL sample+100µL WB Plate method: In a 96-well plate mix 200µL PFS+20µL sample+20µL WB	Konstanz method: In a 96-well plate mix 200µL RPMI+20µL sample+20µL WB PEI Method: In a 96-well plate mix 180µL RPMI + 20µL sample+40µL WB	In a 96-well plate: Mix 50 µL standards/samples+100 µL PFS+50 µL WB	In a 96-well plate: Mix 50 µL standards/samples+100 µL RPMI-C+100 µL PBMCs	In a 96-well plate: Mix 50 µL standards/samples+100 µL RPMI-C+100 µL cells in suspension
	Duration of exposure	10-24 hr		16-24 hr		
	Material used for ELISA	Tube method: centrifuge 2 min @ 10,000 x *g*-test supernatant Plate method: mix each well be pipetting and test resuspended mixture	WB/RPMI/sample mixture	WB/saline/sample mixture	Cell supernatant	Cell supernatant
Known limits of use		Intended for parenteral pharmaceuticals, biological products, and medical devices that have been qualified through interference testing				
Nature of the response assessed		Pyrogenic substances induce the release of proinflammatory cytokines (e.g., IL-1β) from monocytoid cells present in human WB		Pyrogenic substances induce the release of proinflammatory cytokines (e.g., IL-6) from monocytoid cells in WB, PBMC, or immortalized MM6 cells		
Appropriate controls	**Positive control (PC)**	0.5 EU/mL WHO-LPS 94/580 [*E. coli* 0113:h10:K-][5]				
	Negative control (NC)	PFS				
	Positive product control (PPC)	Test substance spiked with endotoxin (0.5 EU/mL or a concentration in middle of standard endotoxin curve)				
	Negative product control (NPC)	Test substance spiked with PFS				
Assay acceptability criteria		• PC OD 1.6-fold>NC OD • PPC OD 1.6-fold>NPC OD • PPC OD should be within 50% to 200% of the PC OD • NC OD≤0.100	• PC OD 1.6-fold>NC OD • PPC OD 1.6-fold>NPC OD • PPC OD should be within 50% to 200% of the PC OD • NC OD≤0.100	• PPC OD should be within 50% to 200% of the PC OD • NC OD<200 pg/mL IL-6 standard	• PPC OD should be within 50% to 200% of the PC OD • 1 EU/mL standard OD>1000 pg/mL IL-6 standard • NC OD<0.15 and NC OD<500 pg/mL IL-6 standard	• PC OD±20% of the expected value (i.e., 0.5 EU/mL) • PPC OD should be within 50% to 200% of the PC OD • NC OD<0.200

	Test Method Component	WB/IL-1β[1]	Cryo WB/IL-1β	WB/IL-6	PBMC/IL-6[1]	MM6/IL-6
Nature of data to be collected and methods used for data collection		• The endotoxin content of a test substance is calculated by comparing the induced IL-1β release with that induced by the endotoxin standard curve concentrations		• The endotoxin content of a test substance is calculated by comparing the induced IL-6 release with that induced by the endotoxin standard curve concentrations		
Type of media in which data are stored		Electronic files				
Exclusion criteria		Mean±SD of the OD for each test substance/standard				
Decision criteria for pyrogenicity		OD TS > OD 0.5 EU/mL EC		EC TS > ELC TS	EC TS > ELC TS[6]	EC TS > ELC TS

Abbreviations: Cryo = Cryopreserved; DMSO = Dimethylsulfoxide; EC = Endotoxin concentration; ELC = Endotoxin limit concentration; EU/mL = Endotoxin units/mL; ELISA = Enzyme-linked immunosorbent assay; HIFCS = Heat-inactivated fetal calf serum; HSA = Human serum albumin; IL = Interleukin; IU = International units; LPS = Lipopolysaccharide; MEM = Minimum essential medium; MM6 = Mono Mac 6; NC = Negative control; NPC = Negative product control; OD = Optical density; PBMC = Peripheral blood mononuclear cells; PC = Positive control; PEI = Paul-Ehrlich-Institut; PFS = Pyrogen free saline; PFW = Pyrogen free water; PPC = Positive product control; SD = Standard deviation; SOP = Standard operating procedure; TS = Test substance; WB = Whole blood; WHO = World Health Organization; x *g* = times gravity

[1]As described in **Section 1.1.1**, a catch-up validation studies were also conducted to evaluate the performance of the WB/IL-1β test method using 96-well plates, and the PBMC/IL-6 test method when using cryopreserved PBMCs. The plating procedure (WB/IL-1β) and the cryopreservation procedure (PBMC/IL-6) are the only differences in the test method protocols (see **Appendix A**).

[2]Medium should be qualified for testing by a valid bacterial endotoxin test (i.e., USP30 NF25<85>) indicating that the endotoxin contamination is <0.06 IU/mL); fetal bovine serum concentration for MM6 cells varies based on whether it is for maintenance/propagation (10%) or assay (2%) conditions.

[3]Dilution of the test material should not exceed the maximum valid dilution (MVD), where MVD = (endotoxin limit concentration)/(detection limit of the assay)

[4]Cell numbers represent viable cells based on trypan blue exclusion

[5]Or another endotoxin calibrated against this standard

[6]Includes a sequential decision strategy in which 3 to 4 donors are tested per substance. 1) If all donors show negative - product is non-pyrogenic; 2) If ≥2 donors show a positive - product is pyrogenic; 3) If only one donor shows a positive, an additional 3 to 4 donors are tested and if no more than one donor is positive (out of 6 to 8 donors) - product is non-pyrogenic; otherwise, product is pyrogenic**.**

[This Page Intentionally Left Blank]

2.2.1 Methods Used to Analyze the Data, Including Methods to Analyze for Interference with the Assay

Once a substance has been tested in the requisite number of donor samples (see **Section 2.2.2**), the resulting sample test medium (as indicated in **Table 2-1**) is assayed in quadruplicate in the relevant cytokine ELISA. Outliers are identified using the nonparametric Dixon's test ($p > 0.05$) (Dixon 1950; Barnett and Lewis 1984), the Grubbs' test (Grubbs 1969) for normally distributed samples, or other statistically acceptable methods (Martin and Roberts 2006) and are excluded from the calculations of endotoxin content (see also **Section 5.3** and **Appendix C**). Endotoxin standard curves are included in each assay, from which the endotoxin content of each replicate is estimated using a 4-parameter logistic model.

As indicated in **Table 2-1**, mean optical density (OD) readings are calculated for the positive and negative control samples, as well as for the relevant positive and negative product controls. The acceptable range of the positive product control (50% to 200% of the positive control response) defines the threshold for interference with the test system. If the positive product control response falls outside of this range, the samples are then assayed at the lowest dilution that does not cause interference.

2.2.2 Decision Criteria and the Basis for the Prediction Model Used to Identify a Pyrogenic Substance

As described in **Section 4.2**, historical RPT data were used to establish a threshold pyrogen dose (i.e., the endotoxin dose at which fever was induced in 50% of the rabbits), which was determined to be 5 EU/kg. Based on the largest allowable volume for injection in rabbits (10 mL/kg), the limit of detection that the *in vitro* pyrogen tests must meet was defined as 0.5 EU/mL. Accordingly, the prediction model for each test method was established based on this limit of detection (i.e., a substance is considered pyrogenic if the mean response is greater than or equal to the 0.5 EU/mL standard).

For three of the test methods, results from multiple donors (Cryo WB/IL-1β [n=5], WB/IL-6 [n=3], and PBMC/IL-6 [n=3 to 4]) are required to determine the potential pyrogenicity of a test substance. In contrast, a single donor sample is used for the WB/IL-1β test method, as is a single cell sample for the MM6/IL-6 test method. As outlined in **Table 2-2**, unlike the Cryo WB/IL-1β test method, the WB/IL-6 and PBMC/IL-6 test methods employ a decision strategy that takes into account the individual responses of each donor sample.

Table 2-2 Prediction Model Used for *In Vitro* Pyrogen Test Methods

Test Method	No. Donors	No. Positive	No. Negative	Decision
PBMC/IL-6	4[1]	4	0	Pyrogenic
		3	1	Pyrogenic
		2	2	Pyrogenic
		1	3	Non-pyrogenic
		0	4	Non-pyrogenic
	3[1]	3	0	Pyrogenic
		2	1	Pyrogenic
		1	2	Non-pyrogenic
		0	3	Non-pyrogenic
WB/IL-6	3	3	3	Pyrogenic
		2	1	Pyrogenic
		1	2	Non-pyrogenic
		0	3	Non-pyrogenic
Cryo WB/IL-1β	5 (pooled)[2]	1	0	Pyrogenic
		0	1	Non-pyrogenic
WB/IL-1β	1	1	0	Pyrogenic
		0	1	Non-pyrogenic
MM6/IL-6	NA[3]	1	0	Pyrogenic
		0	1	Non-pyrogenic

Abbreviations: Cryo = Cryopreserved; IL = Interleukin; MM6 = Mono Mac 6; NA = Not applicable; PBMC = Peripheral blood mononuclear cells; WB = Whole blood

[1]Samples are collected from four donors for the PBMC/IL-6 test method. One donor sample may be excluded based on quality criteria, in which case the prediction model may be applied to results from three donors.

[2]Samples are collected from five donors for the Cryo WB/IL-1β test method and pooled prior to cryopreservation.

[3]Not applicable, because source material is obtained from an immortalized cell line.

2.2.3 Information and Data to be Included in the Study Report and Availability of Standard Forms for Data Collection and Submission

The test report should include the following information, if relevant to the conduct of the study:

Test Substances and Control Substances

- Name and type (e.g., pharmaceutical, biological product, medical device eluate, etc.) of test product
- Purity and composition of the test substance or preparation
- Physicochemical properties, such as physical state, volatility, pH, stability, chemical class, water solubility, relevant to the conduct of the study
- Quality assurance (QA) data and known biological properties
- Treatment of the test/control substances prior to testing, if applicable (e.g., vortexing, sonication, warming; solvent used)
- Stability, if known

Justification of the Specific Protocol(s) Used

Test Method Integrity

- The procedure used to ensure the integrity (i.e., accuracy and reliability) of the test method over time
- If the test method employs proprietary components, documentation of the procedure used to ensure their integrity from lot-to-lot and over time

Criteria for an Acceptable Test

- Acceptable concurrent positive control ranges based on historical data from the testing laboratory (which should be included in the report)
- Acceptable negative control data, including historical control ranges from the testing laboratory (which should be included in the report)

Test Conditions

- Cell system used; donor information, if relevant
- Calibration information for the equipment used for measuring cytokine release (e.g., spectrophotometer)
- Details of test procedure used
- Description of modifications of the test procedure made by the testing laboratory for the substance being tested
- Reference to the laboratory's historical data for the cell system and protocol
- Description of data and QA evaluation criteria used

Results

- Tabulation of data from individual test samples

Description of Other Effects Observed

Discussion of the Results

Conclusion

A Good Laboratory Practice (GLP) QA Statement

- This statement addresses all GLP inspections and audits made during the study, and the dates the results were reported to the Study Director. This statement also serves to confirm that the final report reflects the raw data.

Reporting requirements for GLP-compliant studies are provided in the relevant guidelines (e.g., Organisation for Economic Co-operation and Development [OECD] 1998; U.S. Environmental Protection Agency [EPA] 2003a, 2003b; FDA 2003).

2.3 Basis for Selection of the Test Method Systems

One of the difficulties associated with the currently required pyrogen test methods (i.e., BET and RPT) is that both require extrapolation of the response from a non-human system to the human. In contrast, and as discussed in **Section 1.1.1**, all five of these test methods employ human cells in an attempt to mimic the human fever response *in vitro*.

The WB test methods (i.e., Cryo WB/IL-1β, WB/IL-1β, WB/IL-6) offer the convenience of performing the assay directly on a human blood sample, with minimal pretest preparation. The Cryo WB/IL-1β test method was developed to offer the convenience of an increased time interval between the time of blood collection and the time a test is initiated (since the fresh blood methods require testing within four hr of collection), as well as increased standardization through the pooling of five donor samples to produce a larger sample bank of cells to use in the test. The MM6/IL-6 test method provides increased standardization by using an immortalized cell line that may be maintained in the laboratory indefinitely, and transferred among laboratories. Finally, the PBMC/IL-6 test method was developed in an attempt to improve pyrogen detection sensitivity by using the monocyte fraction of WB, which is considered to be the most sensitive human blood cell type to the presence of endotoxin.

Additional information on standardization of the cellular components required for the test methods is presented in the ECVAM response to ICCVAM PWG questions (see question #5 in **Appendix B**).

2.4 Proprietary Components

Data from the test methods that use the IL-6 endpoint (i.e., WB/IL-6, PBMC/IL-6, MM6/IL-6) were obtained using either an in-house IL-6 ELISA developed by Novartis Pharma AG (Basel, Switzerland) or the Central Laboratory for the Blood Transfusion Service (CLB) Human IL-6 ELISA kit (Amsterdam, The Netherlands). In the ECVAM response to ICCVAM PWG questions (**Appendix B**), it was stated that both IL-6 ELISAs use the same monoclonal anti-IL-6 antibody for detection. At the present time, the Novartis IL-6 ELISA is not available for purchase; however, the CLB IL-6 ELISA kit is commercially available. Importantly, other commercially available IL-6 ELISAs may be individually validated and used in these procedures.

The MM6 cell line was generated by Professor Ziegler-Heitbrock at the University of Munich (refer to **Appendix A**, as well as Section 2.4 of the ECVAM MM6/IL-6 BRD). These cells are currently available from The German Collection of Microorganisms and Cell Cultures (DSMZ). However, a legal agreement with Professor Ziegler-Heitbrock stating that the MM6 cell line will be used for research purposes only is required prior to purchasing the cells. At the present time, any organization (e.g., pharmaceutical company) wishing to use the cells for product testing has to negotiate a fee for provision of the cells and a royalty payment per batch of product tested.

According to the U.S. Patent and Trademark Office (USPTO), patents are held for "Test for determining pyrogenic effect of a material" (U.S. 5,891,728, April 6, 1999), and "Pyrogenicity test for use with automated immunoassay systems" (U.S. 6,696,261 B2, February 24, 2004). These patents cover the WB/IL-1β and WB/IL-6 test methods, respectively. In addition, and related to the WB/IL-1β test method, there is a patent application pending for "Test procedure with biological system - Preparations containing deep-frozen blood are used for determining blood response" (USPTO 436518000).

There are several measures in the study validity criteria that may be used to verify the integrity of proprietary components. As outlined in **Table 2-1**, an endotoxin standard curve is established for each assay, which is in turn used to define the endotoxin activity of the test

substances. In addition, positive and negative controls, along with positive and negative product controls, are used for interference testing, and serve as internal controls for each assay.

2.5 Number of Replicates

2.5.1 Number of Donors

There is no rationale provided for the number of donors included for each test method. As described in **Section 2.2.2**, samples from multiple donors are required for three of the test methods. The Cryo WB/IL-1β test method uses pooled blood from five different donors and the WB/IL-6 and PBMC/IL-6 test methods use blood from at least three donors, which are tested individually. In contrast, a single donor sample is used for the WB/IL-1β test method.

2.5.2 Number of Assay Replicates

Once each substance has been tested in the requisite number of donor samples (see **Section 2.2.2**), the resulting sample test medium is assayed in quadruplicate in the relevant cytokine ELISA. As indicated in **Section 2.2.1**, the nonparametric Dixon's test (Dixon 1950; Barnett and Lewis 1984) or Grubbs' test (Grubbs 1969) for normally distributed samples is used to detect outliers among the replicates. Section 2.5 of the ECVAM BRDs states that four replicates were chosen, as it is considered the minimum number for inclusion in Dixon's test.

2.6 Modifications to the Test Method Protocols Based on ECVAM Validation Study Results

In the MM6/IL-6 test method, prevalidation studies demonstrated that pre-incubation of the cells at a defined initial concentration of 2 x 10^7 cells/50 mL RPMI-C for 24 hr greatly improved test method performance. Therefore, this modification was included in the validation study, and subsequently carried forward to the recommended MM6/IL-6 test method protocol.

For the PBMC/IL-6 test method, a single blood donor was initially used as a source of PBMCs. However, the use of PBMCs from four separate donors (assayed individually) was shown to reduce variability, and this modification was carried forward in the recommended PBMC/IL-6 test method protocol.

No modifications were made to the WB/IL-1β, Cryo WB/IL-1β, and WB/IL-6 test method protocols as a result of the prevalidation or validation testing experiences.

2.7 Differences Between Comparable Validated Test Methods with Established Performance Standards

The differences between the *in vitro* pyrogen test methods and the currently accepted pyrogen test methods (i.e., BET and RPT) are described in **Sections 1.2.3** and **1.3.2**.

[This Page Intentionally Left Blank]

3.0 Substances Used for the Validation of *In Vitro* Pyrogen Test Methods

3.1 Rationale for the Substances or Products Selected for Testing

A validation study should evaluate an adequate subset of substances and product types that are to be tested by the proposed test method. In response to a request for additional information, the rationale for the specific test substances selected for inclusion in the validation studies was provided by ECVAM, which included stability of the endotoxin-spike, relevance, availability/feasibility, and cost (see **Appendix C**). Briefly, to maintain the desired concentration of the endotoxin-spike solution over the time period needed for the validation studies, the test substances and the endotoxin-spike solution were provided separately to the test laboratories and mixed prior to testing. As for relevance, only substances intended for i.v. injection were selected. In addition, test substances consisted solely of marketed parenteral pharmaceuticals that were labeled as free from detectable pyrogens such that these data were available for comparison to the validation study results.

3.2 Number of Substances

A total of 13 substances were included in the performance analysis of each of the five *in vitro* test methods. Ten substances, each spiked with four different concentrations of endotoxin (0, 0.25, 0.5, and 1.0 EU/mL, with 0.5 EU/mL tested in duplicate), were used to evaluate accuracy. Three substances, each spiked with three concentrations of endotoxin (0, 0.5, and 1.0 EU/mL, with 0 EU/mL tested in duplicate), were used to assess intralaboratory reproducibility.

3.3 Identification and Description of Substances Tested

As indicated in **Section 3.1**, the test substances selected for use in the validation studies were marketed parenteral pharmaceuticals. **Table 3-1** lists the 10 test substances used to evaluate accuracy, and **Table 3-2** lists the three test substances used to evaluate reproducibility. In response to a request for additional information, ECVAM provided the lot numbers of the substances used in accuracy evaluation for the validation study, which demonstrated that they were identical (**Appendix C**). However, some of the lots tested in the catch-up validation study for the Cryo WB/IL-1β test method were different (i.e., Fenistil and Sostril) because the original lots were no longer available. One test substance (i.e., Orasthin) was no longer available and was replaced with Syntocinon, which contains the same active ingredient.

Table 3-1 Parenteral Drugs Used in the Validation Studies for Determining Test Method Accuracy[1]

Test Substance[2]	Source	Lot Number(s)	Active Ingredient	Indication	MVD (-fold)
Beloc®	Astra Zeneca	DA419A1	Metoprolol tartrate	Heart dysfunction	140
Binotal®	Grünenthal	117EL2	Ampicillin	Antibiotic	140
Ethanol 95%	B. Braun	2465Z01	Ethanol	Diluent	35
Fenistil®	Novartis	21402 26803[3]	Dimetindenmaleat	Antiallergic	175
Glucose 5%	Eifelfango	1162 3132[3]	Glucose	Nutrition	70
MCP®	Hexal	21JX22	Metoclopramid	Antiemetic	350
Orasthin®	Hoechst	W015	Oxytocin	Initiation of delivery	700
Sostril®	Glaxo Wellcome	1L585B 3H01N[3]	Ranitidine	Antiacidic	140
Syntocinon®	Novartis	S00400	Oxytocin	Induction of labor	-
Drug A - 0.9%NaCl	-	-	0.9% NaCl	-	35
Drug B - 0.9% NaCl	-	-	0.9% NaCl	-	70

Abbreviations: MVD = Maximum valid dilution
[1]Each substance was tested in all five *in vitro* pyrogen test methods.
[2]Each test substance was spiked with 0, 0.25, 0.5, or 1.0 endotoxin units/mL (EU/mL) of endotoxin (WHO-LPS 94/580 [*E. coli* O113:H10:K-]), with 0.5 EU/mL tested in duplicate. Each sample contained the appropriate spike concentration when tested at its MVD.
[3]Indicates the lot number used in the catch-up validation study for the Cryopreserved Whole Blood/Interleukin-1β test method.

Table 3-2 Parenteral Drugs Used in the Validation Studies for Determining Test Method Reproducibility[1]

Test Substance[2]	Source	Agent	Indication
Gelafundin®	Braun Melsungen	Gelatin	Transfusion
Haemate®	Aventis	Factor VIII	Hemophilia
Jonosteril®	Fresenius	Electrolytes	Infusion

[1]Each substance was tested in all five *in vitro* pyrogen test methods.
[2]Each test substance was spiked with 0, 0.5, or 1.0 endotoxin units/mL (EU/mL) of endotoxin (WHO-LPS 94/580 [*E. coli* O113:H10:K-]), with 0 EU/mL tested in duplicate. Each sample contained the appropriate spike concentration when tested at its maximum valid dilution.

3.4 Sample Coding Procedure

According to the ECVAM BRDs (Section 3.4), the 10 test substances and the four spike concentrations used for the evaluation of accuracy were blinded to the testing laboratories. For the reproducibility analyses, although the three spike concentrations were blinded to the participating laboratories, the identities of the three test substances were not.

3.5 Rationale for the Selection of the Recommended Reference Substances

Reference substances are used to assess the accuracy and reliability of a proposed, mechanistically and functionally similar test method and are a representative subset of those used to demonstrate the reliability and accuracy of the validated reference test method (in this case, the RPT). These substances should:

- Represent the range of responses that the validated test method is capable of measuring or predicting
- Have produced consistent results in the validated test method
- Produce responses that reflect the accuracy of the validated test method
- Have well-defined chemical structures and/or compositions
- Be readily available
- Not be associated with excessive hazard or prohibitive disposal costs

For evaluating test method performance, each of the test substances used in the ECVAM validation studies was spiked with a Gram-negative endotoxin standard (WHO-LPS 94/580 [*E. coli* O113:H10:K-]). Two different sources of endotoxin (i.e., *E. coli* EC-5 and *E. coli* EC-6), which were reported to be identical to the WHO standard, were used in the validation studies (Hochstein et al. 1994; Hoffman et al. 2005a). Endotoxin was selected as a "model" pyrogen for inclusion based on its availability in a standardized form and because of the known ability of monocytic cells to respond to endotoxin-based pyrogens. Endotoxin was also used as a positive control and for qualifying the *in vitro* test methods during interference testing. It is also used when performing the BET. As described in **Section 4.0**, the response of the reference test method (i.e., RPT) to endotoxin is well documented. For this reason, the threshold pyrogen dose used for establishing the decision criteria for the *in vitro* test methods was based on historical RPT data. Importantly, no other non-endotoxin-based pyrogenic substances are presently available in a standardized form.

[This Page Intentionally Left Blank]

4.0 *In Vivo* Reference Data for the Assessment of Test Method Accuracy

4.1 Description of the Protocol Used to Generate *In Vivo* Data

4.1.1 The Rabbit Pyrogen Test

The RPT protocols most widely accepted by regulatory agencies are outlined in the USP (USP 2007b), the U.S. Code of Federal Regulations (FDA 2005), the European Pharmacopeia ([EP], EP 2005a), and the Japanese Pharmacopeia ([JP], JP 2001), and are summarized in **Table 4-1**. The RPT involves measuring the temperature increase in rabbits following an i.v. injection (via the ear vein) of a test substance in a dose not to exceed 10 mL/kg injected within a period of not more than 10 min. Initially, three rabbits are injected and the increase (or decrease) in temperature relative to the baseline value is measured at 30-min intervals for up to three hr. The resulting data are used to calculate an overall temperature increase by adding the results from all three animals, which is then used to assign a label of pyrogenic or non-pyrogenic.

[This Page Intentionally Left Blank]

Table 4-1 Test Guidelines for the Rabbit Pyrogen Test

RPT Protocol Component	Reference			
	21 CFR 610.13 (FDA 2005)	**EP5.0 2.6.8 (EP 2005a)**	**JP XIV (JP 2001)**	**USP30 NF25 <151> (USP 2007b)**
Number of rabbits	3 or 8[1]	3, 6, 9, or 12[1]	3 or 8[1]	3 or 8[1]
Rabbit species/strain	Not specified	Not specified	Not specified	Not specified
Exclusion criteria for rabbits during the initial selection of rabbits	• Used in a negative pyrogen test in the preceding 2 days • Used in a pyrogen test in which its temperature rose ≥0.6°C in the preceding 2 weeks	• Weight<1.5 kg • Decreased weight in the preceding week • Used in a negative pyrogen test in the preceding 3 days • Used in a positive pyrogen test in the preceding 3 weeks	• Weight<1.5 kg • Decreased weight in the preceding week • Previously used in a positive pyrogen test • Rabbits from negative pyrogen tests may be reused only when a "as a long a resting period as possible is taken"	• Used in a negative pyrogen test in the preceding 2 days • Used in a pyrogen test in which its temperature rose ≥0.6°C in the preceding 2 weeks
Testing room conditions	20 to 23°C	Within 3°C of the housing quarters (temperature not specified)	20 to 27°C and constant humidity	20 to 23°C
Food/water during test	Food withheld during the test, but water available at all times	Food withheld overnight and until end of the test. Water withheld during the test.	Food withheld beginning several hrs. prior to first temperature recording and until the end of the test.	Food withheld during the test period, but water available at all times
Depth of temperature probe in rectum	Not less than 7.5 cm	Approximately 5 cm	6-9 cm	Not less than 7.5 cm
Preliminary test	≤7 days prior to main test, perform all procedures used for the main test except the injection.	• 1-3 days prior to main test, treat test animals with an injection of warmed (38.5°C) pyrogen-free saline • Record temperature at 90 min prior to injection and every 30 min thereafter up to 3 hr. • Exclude any rabbits with an increase of >0.6°C	Not specified	≤7 days prior to main test, perform all procedures used for the main test except the injection.
Baseline temperature	• Record temperature ≤ 30 min prior to injection • For any group of rabbits, use only if baseline temperatures do not vary>1°C among rabbits • Exclude rabbits with baseline temperature>39.8°C	• Mean of two temperature recordings at 40 min and 10 min prior to injection • Exclude rabbits if variation >0.2°C between measurements noted • Exclude rabbits with initial temperature >39.8°C or <38.0°C	• Record temperature three times at one-hr intervals prior to injection • Assuming no appreciable variability among recordings, use the last recording as the baseline value. • Exclude animals if 2nd and 3rd temperature measurements exceed 39.8°C	• Record temperature ≤30 min prior to injection • For any group of rabbits, use only if baseline temperatures do not vary >1°C among rabbits • Exclude rabbits with baseline >39.8°C
Injection volume	≥3 mL/kg *BUT* ≤10mL/kg	≥0.5 mL/kg *BUT* ≤10mL/kg	10 mL/kg, unless otherwise specified	≤10 mL/kg

Injection time	≤10 min	≤4 min, unless otherwise indicated	Not specified, but injection should occur within 15 min of the third pretest temperature recording	≤10 min
Injection site	Marginal ear vein	Marginal ear vein	Marginal ear vein	Marginal ear vein
Pre-warming of test material	37°C±2°C	38.5°C	37°C	37°C±2°C
Temperature recording intervals after injection	30 min intervals for 1 to 3 hr	≤30 min intervals for 3 hr	1 hr intervals for 3 hr	30 min intervals for 1 to 3 hr

Abbreviations: CFR = U.S. Code of Federal Regulations; EP = European Pharmacopeia; FDA = U.S. Food and Drug Administration; JP = Japanese Pharmacopeia; RPT = Rabbit pyrogen test; USP = United States Pharmacopeia

[1]Each test is initially conducted with three animals and additional animals are tested to resolve equivocal results in the first three animals

4.1.2 Current In Vivo Pyrogen Test Method Protocols

As indicated in **Table 4-1**, U.S. and international regulatory agencies have tailored the RPT protocol to suit their specific needs and goals in protecting human health. The current test method protocols (i.e., FDA 2005; EP 2005a; JP 2001; USP 2007b) recommend using healthy, adult rabbits with no specific breed/strain requirements. Rabbits are to be adequately acclimated to their surroundings and housed in an environment free from excessive external stimuli. Each rabbit is conditioned prior to the test with a sham test that includes all of the procedural steps except the injection (see also **Section 1.2**). Reuse of test rabbits is permitted only after an appropriate withdrawal period has been completed (see also **Section 1.2**).

The test is conducted in a room that is designated solely for pyrogen testing, in which the temperature is within 3°C of the uniform temperature of the housing room (i.e., 20°C±3°C). Food is withheld during the test, but access to water is continuous. The baseline temperature, which is used to calculate the increase in temperature during the test, is measured 30-40 min prior to injection of the test substance. In each group of rabbits tested, the variation in baseline temperature among the rabbits should not vary more than 1°C, and rabbits with an initial temperature greater than 39.8°C are excluded from testing.

The test substance is pre-warmed to approximately 37°C and injected (≤10 mL/kg) into the marginal ear vein, completing each injection within 10 min. The rectal temperature is recorded at 30-min intervals for up to three hr after the injection. The decision criteria outlined in **Table 4-2** are then used to determine a pyrogenic response. As shown in **Table 4-2**, the decision criteria by which labels of pyrogenic or non-pyrogenic are assigned vary among the USP, FDA, EP, and JP test guidelines.

Table 4-2 Decision Criteria for Determining a Pyrogenic Response in the Rabbit Pyrogen Test

RPT Protocol	No. Rabbits	Product passes if:	Product fails if:
USP30 NF25<151> (USP 2007b)	3	0/3 rabbits show an increase of ≥0.5°C	NA[1]
	5[1]	≤3/8 rabbits show an increase of ≥0.5°C *AND* the summed responses ≤3.3°C	>3/8 rabbits show an increase of ≥0.5°C *AND/OR* the sum of all responses >3.3°C
21 CFR 610.13 (FDA 2005)	3	0/3 rabbits show an increase of ≥0.5°C	NA[1]
	5[1]	≤3/8 rabbits show an increase of ≥0.6°C *AND* the summed responses ≤3.7°C	>3/8 rabbits show an increase of ≥0.6°C *AND/OR* the summed responses >3.7°C
EP5.0 2.6.8 (EP 2005a)	3	Summed responses ≤1.15°C	Summed responses >2.65°C
	6[2]	Summed responses ≤2.80°C	Summed responses >4.30°C
	9[2]	Summed responses ≤4.45°C	Summed responses >5.95°C
	12	Summed responses ≤6.60°C	Summed responses >6.60°C
JP XIV (JP 2001)	3	3/3 rabbits show an increase of <0.6°C *AND* the summed responses ≤1.4°C	≥2/3 rabbits show an increase ≥0.6°C
	5[3]	≥4/5 rabbits show an increase <0.6°C	≥2/5 rabbits show an increase ≥0.6°C

CFR = U.S. Code of Federal Regulations; EP = European Pharmacopeia; FDA = U.S. Food and Drug Administration; JP = Japanese Pharmacopeia; NA = Not applicable; USP = United States Pharmacopeia; RPT = Rabbit pyrogen test

[1]If ≥1/3 rabbits show an increase of ≥0.5°C, continue test with an additional five rabbits.

[2]Three additional animals are tested when the summed responses falls in between the previous range.

[3]Five additional animals are tested when neither criterion is met, and results are based on these five animals only.

4.2 Reference Data Used to Assess *In Vitro* Test Method Accuracy

The ECVAM BRDs state that due to ethical and legal reasons, the RPT was not conducted in parallel to the *in vitro* test methods. Instead, historical RPT data produced over a 5-year period at the Paul-Ehrlich Institut (PEI), which is the German Federal Agency of Sera and Vaccines, were used (Hoffmann et al. 2005a). These data were generated for internal quality control studies from 171 rabbits (Chinchilla Bastards). Chinchilla Bastards are reported to be a more sensitive strain than the New Zealand White rabbit strain for pyrogenicity testing (Hoffmann et al. 2005b). However, neither the USP (USP 2007b) nor the EP (EP 2005a) prescribes a specific rabbit strain for the RPT.

4.3 Availability of Original Records for the *In Vivo* Reference Data

Section 4.1 of each ECVAM BRD indicates that the PEI provided the historical RPT data.

4.4 *In Vivo* Data Quality

The historical RPT studies were conducted at the PEI, which supports regional German regulatory authorities, provides marketing approval of certain marketed biological products (e.g., sera, vaccines, test allergens), and functions as a WHO collaborating center for QA of blood products and *in vitro* diagnostics. The unit for pyrogen and endotoxin testing of the PEI is accredited following ISO/IEC 17025 (International Standards Organization [ISO]

2005). In a request for additional information from ECVAM, it was stated that the RPT data was generated according to the EP monograph, but the detailed protocol used by this laboratory was not provided.

4.5 Availability and Use of Toxicity Information from the Species of Interest

A number of studies have concluded that humans and rabbits have approximately the same threshold to pyrogenic stimulation, although higher doses are more pyrogenic and more toxic in humans (Co Tui and Schrift 1942; Westphal 1956; Keene et al. 1961). Moreover, Greisman and Hornick (1969) compared three purified endotoxin preparations in rabbits and in male volunteers and showed that the threshold pyrogenic dose was similar in both species. However, the dose-response relationships for humans were considerably steeper than those for the rabbit at each dose tested.

As stated in **Section 1.2.1**, the major regulatory requirement for pyrogenicity testing is for end-product release of human and animal parenteral drugs, medical devices, and human biological products. The results from such testing are used to limit, to an acceptable level, the risks of febrile reactions from injection and/or implantation of the product of concern.

Endotoxin can produce a number of acute effects on human health. McKinney et al. (2006) reported increased cytokine expression patterns in a cohort of subjects experiencing systemic adverse events (i.e., fever, rash, lymphadenopathy) after smallpox vaccine administration. Martich et al. (1993) studied systemic, cardiovascular, pulmonary, cytokine release, and the inflammatory response resulting from i.v. injection of small doses of endotoxin in humans to understand mechanisms of sepsis and septic shock. Burrell (1994) later reviewed the available literature on the adverse human responses to bacterial endotoxin. In addition, environmental or chronic exposure to inhaled bacterial endotoxin (present in soil, in water, and on vegetation) may lead to an inflammation in the airways and/or gastrointestinal disturbances (Rylander 2002). Therefore, for protection of both human and animal health, it is vital that the test method employed provide an accurate estimation of the potential for a pyrogenic reaction.

4.6 Information on the Accuracy and Reliability of the *In Vivo* Test Method

Hoffmann et al. (2005a) modeled the sensitivity and specificity of the RPT using historical data (summarized in **Section 4.2**) to establish a threshold pyrogen dose (i.e., the endotoxin dose at which fever was induced in 50% of the rabbits). A threshold value of 0.5 EU/mL was defined by regression analysis of the data. The performance characteristics of the RPT (i.e., sensitivity and specificity) were then determined using a 2 x 2 contingency table, incorporating the parameters obtained from the regression analysis. The authors considered the prevalence of the endotoxin spikes included in the ECVAM accuracy evaluations in the validation studies (i.e., 0 EU/mL: 20%; 0.25 EU/mL: 20%; 0.5 EU/mL: 40%; 1.0 EU/mL: 20%) and applied the threshold pyrogen dose of 0.5 EU/mL to calculate theoretical values for sensitivity (58%) and specificity (88%) of the RPT.

The accuracy and reliability of the RPT for endotoxin testing has been considered adequate for U.S. and international regulatory needs for many years. Since its inclusion in the USP in 1941, the RPT has been used extensively and is the preferred method for detection of pyrogenicity for product development, because of the inability of the BET to detect non-endotoxin pyrogens.

[This Page Intentionally Left Blank]

5.0 Test Method Data and Results

5.1 Test Method Protocol

The Standard Operating Procedures (SOPs) used during the ECVAM validation studies are included in **Appendix A**. As described in **Section 2.1**, there are many similarities among the protocols for each of the *in vitro* pyrogen test methods, with very few notable differences other than the type of cells used (i.e., WB cells, PBMCs, monocytoid cell line) and the proinflammatory cytokine assayed (i.e., IL-1β or IL-6). These similarities and differences are outlined in **Table 2-1**. An internationally accepted endotoxin standard (i.e., WHO-LPS 94/580 [*E. coli* 0113:h10:K-]) was used to spike samples of saline or marketed parenteral pharmaceuticals. The same pharmaceuticals were used to create the spiked samples for all five test methods (see **Table 3-1** and **3-2**). These samples were included in a series of studies designed to determine the relevance and reliability of each of the *in vitro* pyrogen test methods.

5.2 Availability of Copies of Original Data Used to Evaluate Test Method Performance

ECVAM provided raw data from the validation studies in an electronic format (Excel® spreadsheets) that consisted of OD_{450} measurements for all replicates included in each of the validation studies.

NICEATM attempted to obtain additional *in vitro* and/or *in vivo* pyrogen test method data. A *Federal Register* (*FR*) notice (Vol. 70, No. 241, pp. 74833-74834, December 16, 2005) was published requesting original *in vitro* pyrogen test method and reference data from the currently used pyrogen test methods (i.e., RPT and/or BET). In addition, the *FR* notice was sent directly to more than 100 interested stakeholders internationally. Despite these efforts, no additional data were submitted.

5.3 Description of the Statistical Approaches Used to Evaluate the Resulting Data

Details of the statistical approaches used to evaluate the accuracy and reliability of each of the five *in vitro* test methods are included in Section 5.3 of each ECVAM BRD. Briefly, as indicated in **Section 3.2**, 10 substances (each spiked with four concentrations of endotoxin, with one concentration spiked in duplicate) were tested in each test method to evaluate accuracy, while three substances (each spiked with three concentrations of endotoxin, with one tested in duplicate) were used to evaluate test method reproducibility. Varying concentrations of endotoxin-spiked saline were tested for the analysis of intralaboratory repeatability.

The evaluation of intralaboratory repeatability included coefficient of variation (CV) analysis of the log-transformed OD_{450} measurements for the replicates of each endotoxin concentration. Boxplots were also generated to demonstrate variability among these values for each concentration. Similar analyses were conducted for the three substances used to assess intra- and inter-laboratory reproducibility.

The reproducibility analysis incorporated the decision criteria that were developed to differentiate between pyrogenic and non-pyrogenic materials (using a threshold value of 0.5

EU/mL). In all reproducibility analyses, a single run consisted of each of the substances (as described above and in **Section 3.2**) assayed in quadruplicate. Acceptability criteria for each run included a CV analysis to remove highly variable samples from the analyses. This criterion ranged from a CV<0.25 to <0.45, depending on the test method being considered. For the measurement of intralaboratory reproducibility, pair-wise comparisons between the runs were determined and the associations between runs expressed as a percentage of agreement between two individual laboratories. It should be noted that this analysis takes into account the agreement of the resulting pyrogenicity decision (i.e., pyrogenic or non-pyrogenic), but does not consider whether the decision is correct. The correlations (expressed as a percentage of agreement) between pairs of the independent runs (i.e., run 1 vs. run 2; run 1 vs. run 3; run 2 vs. run 3) were determined and the mean of these three values was calculated. Similar analyses were conducted for an assessment of interlaboratory reproducibility, in which pairwise comparisons between laboratories were determined and the associations were expressed as a percentage of agreement. This analysis included each run from each laboratory (n=3 per laboratory) and all possible interlaboratory combinations were compared. Similar to the intralaboratory analysis, this analysis takes the resulting pyrogenicity call from each run in each laboratory into consideration, but does not consider whether the call is correct. **Section 7.0** provides additional details and the resulting data from these analyses.

For the accuracy analysis, 2 x 2 contingency tables were constructed using the decision criteria defined in **Table 4-2** to assign a pyrogenicity call. Each run for each sample from each laboratory was considered independently. Accordingly, the *in vitro* call was compared to the "true status" (based on the known endotoxin spike concentration) of the sample. The resulting accuracy statistics were calculated based on the overall database for each test method. Similar to the reproducibility analyses, acceptability criteria for each run included a CV analysis to remove highly variable samples from the analyses, for which a range of CV<0.25 to CV<0.45 was used, depending on the test method being considered. **Section 6.0** provides additional details and the resulting data from these analyses.

Outliers were identified and eliminated using a two-step procedure. In the first step, replicates with an extremely large variation were identified by comparing the CV for the replicates with the extracted maximal CV (CV_{max}). If the CV for the replicates was smaller than the CV_{max}, then the data were analyzed without modification. However, if the replicates failed to pass this initial test, then the data were transformed with the natural logarithm and examined for outliers using the nonparametric Dixon's test (Dixon 1950; Barnett and Lewis 1984) or the Grubbs' test (Grubbs 1969) for normally distributed samples. If one observation was responsible for the large variation, then the observation was excluded. If the variation was due to all observations, then the entire set of replicates was excluded from further analysis. Additional information on the analytical procedure used to identify and eliminate outlier observations can be found in the materials provided by ECVAM (see **Appendix C**).

5.4 Summary of Results

Graphical representations of the repeatability and reproducibility analyses are provided in Section 5.2 of each ECVAM BRD (see **Appendix A**). The tabulated results from which the intra- and inter-laboratory reproducibility analyses and accuracy analyses can be conducted are provided in Section 5.4 of the ECVAM BRDs. The tables in that section include the test

substance name, the endotoxin spike concentration, the pyrogenicity call for each *in vitro* run, and the "true status" of each test substance.

5.5 Use of Coded Chemicals and Compliance with GLP Guidelines

Ideally, all data supporting the validity of a test method should be obtained using coded chemicals and reported in accordance with GLP guidelines (i.e., OECD 1998; EPA 2003a, 2003b; FDA 2003). **Section 3.4** indicates that the 10 test substances and the four spike concentrations used for the accuracy evaluation were blinded to the testing laboratories. However, although the three spike concentrations were blinded to the participating laboratories for the reproducibility studies, the identity of the three test substances was not blinded.

5.6 Lot-to-Lot Consistency of Test Substances

Lot-to-lot consistency of test substances is evaluated to ensure that the same substance, with the same physicochemical properties, is used for the duration of the study. In these studies, the test substances were released from clinical lots of parenteral pharmaceuticals, which implied that they had been subjected to rigorous chemical manufacturing control analyses to verify that the compositions are consistent. However, the specific lot numbers for the test substances used in the validation study were not initially provided in the ECVAM BRDs. In response to a request for additional information, ECVAM provided this information (**Table 3-1** and **Appendix C**). In addition, the international standard for Gram-negative endotoxin, WHO-LPS 94/580 (*E. coli* O113:H10:K-), was used as the spike solution, which provides a measure of consistency for the positive control substance and the spike substance.

5.7 Availability of Data for External Audit

As described in **Section 8.4**, all records are stored and archived by the participating laboratories and are available for inspection.

[This Page Intentionally Left Blank]

6.0 Relevance of the *In Vitro* Pyrogen Test Methods

6.1 Accuracy of *In Vitro* Pyrogen Test Methods

A critical component of an ICCVAM evaluation of the validation status of a test method is an assessment of its relevance. The measure of relevance used in this evaluation is the performance of the new test in identifying pyrogens as compared to the performance of the current reference method (ICCVAM 2003). This aspect of assay performance is typically evaluated by calculating:

- Accuracy (also referred to as concordance): the proportion of correct outcomes (positive and negative) of a test method
- Sensitivity: the proportion of true positive substances that are correctly classified as positive
- Specificity: the proportion of true negative substances that are correctly classified as negative
- Positive predictivity: the proportion of correct positive responses among substances testing positive
- Negative predictivity: the proportion of correct negative responses among substances testing negative
- False positive rate: the proportion of true negative substances that are falsely identified as positive
- False negative rate: the proportion of true positive substances that are falsely identified as negative

The ability of the *in vitro* pyrogen test methods to correctly identify the presence of Gram-negative endotoxin was evaluated using parenteral pharmaceuticals spiked with endotoxin (WHO-LPS 94/580 [*E. coli* O113:H10:K-]). As described in **Section 3.2**, 10 substances (see **Table 3-1**) spiked with four concentrations of endotoxin (with one concentration in duplicate) were used for the evaluation. The individual spike concentrations in each substance were tested once, using each test method, in three different laboratories, providing a total of 150 runs (i.e., 10 substances x 5 spike solutions x 3 laboratories = 150). The quality criteria outlined in **Table 2-1** were used to identify outliers. These outliers were subsequently excluded from the evaluation, which resulted in less than a total of 150 runs per evaluation.

As described in **Section 4.2**, no RPTs were conducted in parallel with the *in vitro* pyrogen test methods during the ECVAM validation studies. Instead, historical RPT data from rabbits tested with endotoxin were used to establish a threshold pyrogen dose (i.e., the endotoxin dose at which fever was induced in 50% of the rabbits). This historical data were subsequently used to establish the limit of detection (i.e., 0.5 EU/mL) that the *in vitro* test methods being validated must meet. Accordingly, the *in vitro* call was compared to the "true status" (based on the known endotoxin spike concentration) of the sample. The resulting calls were used to construct 2x2 contingency tables, which were used to calculate the resulting test method performance values.

6.1.1 Relevance of the Cryo WB/IL-1β Test Method

Of the 150 available runs for the Cryo WB/IL-1β test method, 10 runs showed excessive variability but no significant outliers among the four replicates (i.e., CV >45%) resulting in their exclusion from the analysis. An additional 20 runs (from one of the three participating laboratories) did not qualify according to one or more of the criteria outlined in **Table 2-1**. Therefore, a total of 120 runs were used in the performance analysis which showed that the Cryo WB/IL-1β test method has an accuracy of 92% (110/120), a sensitivity of 97% (75/77), a specificity of 81% (35/43), a false negative rate of 3% (2/77), and a false positive rate of 19% (8/43) (see **Table 6-1**).

Table 6-1 Accuracy of *In Vitro* Pyrogen Test Methods[1]

Test Method	Accuracy[2]	Sensitivity[3]	Specificity[4]	False Negative Rate[5]	False Positive Rate[6]
Cryo WB/IL-1β	92% (110/120)	97% (75/77)	81% (35/43)	3% (2/77)	19% (8/43)
MM6/IL-6	93% (138/148)	96% (85/89)	90% (53/59)	5% (4/89)	10% (6/59)
PBMC/IL-6	93% (140/150)	92% (83/90)	95% (57/60)	8% (7/90)	5% (3/60)
PBMC/IL-6 (Cryo)[7]	87% (130/150)	93% (84/90)	77% (46/60)	7% (6/90)	23% (14/60)
WB/IL-6	92% (136/148)	89% (79/89)	97% (57/59)	11% (10/89)	3% (2/59)
WB/IL-1β (Tube)	81% (119/147)	73% (64/88)	93% (55/59)	27% (24/88)	7% (4/59)
WB/IL-1β (96-well plate)[8]	93% (129/139)	99% (83/84)	84% (46/55)	1% (1/84)	16% (9/55)

Abbreviations: Cryo = Cryopreserved; EU/mL = Endotoxin units per milliliter; IL = Interleukin; MM6 = Mono Mac 6; PBMC = Peripheral blood mononuclear cells; WB = Whole blood

[1]Data shown as a percentage (number of correct runs/total number of runs), based on results of 10 parenteral drugs tested in each of three different laboratories. Samples of each drug were tested with or without being spiked with a Gram-negative endotoxin standard (0, 0.25, 0.5, or 1.0 EU/mL, with 0.5 EU/mL tested in duplicate).

[2]Accuracy = the proportion of correct outcomes (positive and negative) of a test method.

[3]Sensitivity = the proportion of all positive substances that are classified as positive.

[4]Specificity = the proportion of all negative substances that are classified as negative.

[5]False negative rate = the proportion of all positive substances that are falsely identified as negative.

[6]False positive rate = the proportion of all negative substances that are falsely identified as positive.

[7]A modification of the PBMC/IL-6 test method that uses Cryo PBMCs.

[8]A modification of the WB/IL-1β test method that uses 96-well plates instead of tubes for the test substance incubation.

6.1.2 Relevance of the MM6/IL-6 Test Method

Of the 150 available runs for the MM6/IL-6 test method, two showed excessive variability among the four replicates (i.e., CV >25%), resulting in their exclusion from the analysis. No runs were excluded based on the criteria outlined in **Table 2-1**. Therefore, a total of 148 runs was used in the performance analysis. Based on this analysis, the MM6/IL-6 test method has an accuracy of 93% (138/148), a sensitivity of 96% (85/89), a specificity of 90% (53/59), a false negative rate of 4% (4/89), and a false positive rate of 10% (6/59) (see **Table 6-1**).

6.1.3 Relevance of the PBMC/IL-6 Test Method

None of the 150 available runs for the PBMC/IL-6 test method showed excessive variability (i.e., CV >40%) and all runs met the criteria outlined in **Table 2-1**. Therefore, all 150 runs were included in the performance analysis. Based on this analysis, the PBMC/IL-6 test method has an accuracy of 93% (140/150), a sensitivity of 92% (83/90), a specificity of 95% (57/60), a false negative rate of 8% (7/90), and a false positive rate of 5% (3/60) (see **Table 6-1**).

6.1.3.1 Relevance of the PBMC/IL-6 Method When Using Cryo PBMCs

As indicated in **Table 2-1**, the PBMC/IL-6 test method protocol was also conducted using a modified protocol that included Cryo PBMCs. None of the 150 available runs for this modification of the PBMC/IL-6 test method showed excessive variability (i.e., CV >40%) and all runs met the criteria outlined in **Table 2-1**. Therefore, all runs were included in a performance analysis. Based on this analysis, the PBMC/IL-6 test method, when using Cryo PBMCs, has an accuracy of 87% (130/150), a sensitivity of 93% (84/90), a specificity of 77% (46/60), a false negative rate of 7% (6/90), and a false positive rate of 23% (14/60). The high false positive rate can be attributed to a large number of false positives (50% [10/20]) in one of the three laboratories (the false positive rate in the remaining two laboratories is 10%).

6.1.4 Relevance of the WB/IL-6 Test Method

None of the 150 available runs for the WB/IL-6 test method showed excessive variability (i.e., CV >45%) and all runs met the criteria outlined in **Table 2-1**. However, two samples were mishandled by one of the testing laboratories, and thus the two associated runs were excluded from the analysis. As a result, 148 runs were included in the performance analysis for the detection of Gram-negative endotoxin. Based on this analysis, the WB/IL-6 test method has an accuracy of 92% (136/148), a sensitivity of 89% (79/89), a specificity of 97% (57/59), a false negative rate of 11% (10/89), and a false positive rate of 3% (2/59) (see **Table 6-1**).

6.1.5 Relevance of the WB/IL-1β Test Method

Of the 150 available runs for the WB/IL-1β test method, three showed excessive variability among the four replicates (i.e., CV >45%), resulting in their exclusion from the analysis. No runs were excluded based on the criteria outlined in **Table 2-1**. Therefore, a total of 147 runs was used in the performance analysis. Based on this analysis, the WB/IL-1β test method has an accuracy of 81% (119/147), a sensitivity of 73% (64/88), a specificity of 93% (55/59), an false negative rate of 27% (24/88), and a false positive rate of 7% (4/59) (see **Table 6-1**). Improved performance statistics for the WB/IL-1β test method associated with the use of 96-well plates is summarized below (**Section 6.1.5.1**).

6.1.5.1 Relevance of the WB/IL-1β Test Method When Using 96-Well Plates

As indicated in **Table 2-1**, the WB/IL-1β test method protocol was also conducted using a modified protocol that used 96-well plates instead of individual tubes. Of the 150 available runs for this modification of the WB/IL-1β test method, 11 showed excessive variability (i.e., CV >45%). No runs were excluded based on the criteria outlined in **Table 2-1**. Therefore, a

total of 139 runs were included in a performance analysis. Based on this analysis, the WB/IL-1β test method, when using 96-well plates, has an accuracy of 93% (129/139), a sensitivity of 99% (83/84), a specificity of 84% (46/55), a false negative rate of 1% (1/84), and a false positive rate of 16% (9/55).

6.2 Summary of the Performance Statistics for *In Vitro* Pyrogen Test Methods

The performance of the *in vitro* pyrogen test methods for the detection of Gram-negative endotoxin (based on 10 parenteral pharmaceuticals, each spiked with four concentrations of endotoxin, with one spiked in duplicate) was evaluated. As outlined in **Table 6-1**, this analysis indicated that the accuracy among the test methods ranged from 81% to 93%, sensitivity ranged from 89% to 99%, specificity ranged from 81% to 97%, false negative rates ranged from 1% to 27%, and false positive rates ranged from 3% to 23%.

A comparison of the results for the *in vitro* test methods indicates that the number of runs excluded was greatest for the Cryo WB/IL-1β and WB/IL-1β (plate method) test methods, which had 30 and 11 runs excluded, respectively. No other test method had more than three runs excluded.

6.2.1 Discordant Results

It was not possible to make a direct comparison between the RPT and *in vitro* pyrogen test results without the availability of parallel testing data (i.e., same test substance tested using the *in vitro* and *in vivo* methods). Therefore, *in vitro* results that are discordant from the RPT could not be identified with these studies. Discordant results reflect either a failure of the *in vitro* test method to identify Gram-negative endotoxin (i.e., false negative) when spiked into a test substance at 0.5 EU/mL (i.e., the threshold concentration established based on historical data from the RPT) or 1.0 EU/mL, or to incorrectly indicate the presence of Gram-negative endotoxin (i.e., false positive) when spiked into a test substance at 0 or 0.25 EU/mL. As shown in **Table 6-2**, false positive rates ranged from 7% to 47% when spiked into a test substance at 0.25 EU/mL and from 0% to 3% when spiked with 0 EU/mL. Similarly, false negative rates ranged from 2% to 39% when spiked into a test substance at 0.5 EU/mL and from 0% to 3% when spiked with 1.0 EU/mL.

6.2.2 Strengths and Limitations of In Vitro Pyrogen Test Methods

The limitations of these test methods have not been fully explored and identified. As described in **Section 3.0**, the substances tested do not adequately represent the range of products that are tested with these methods. For this reason, pre-testing product specific validation will be necessary to establish if a particular test substance/material is appropriate for evaluation using these *in vitro* test methods. A recognized limitation of the *in vitro* methods is the lack of data to determine their responses to, and suitability for, non-endotoxin pyrogens that can be detected by the RPT. Additional limitations of these test methods are outlined in the ECVAM response to ICCVAM PWG questions (see question #4 in **Appendix B**). However, an advantage to these *in vitro* test methods is that they are derived from human tissues, and thus avoid potential uncertainty associated with cross-species extrapolation.

Table 6-2 Predictivity of *In Vitro* Pyrogen Test Methods for Each Endotoxin Spike Concentration[1]

Test Method	Endotoxin Spike Concentration								Overall Totals	
	Negative for Pyrogen (< 0.5 EU/mL)				Positive for Pyrogen (≥ 0.5 EU/mL)					
	0 EU/mL		0.25 EU/mL		0.5 EU/mL		1.0 EU/mL		False Negative	False Positive
	Correct	False Positive[2]	Correct	False Positive	False Negative[3]	Correct	False Negative	Correct		
Cryo WB/IL-1β	100% (24/24)	0% (0/24)	58% (11/19)	42% (8/19)	4% (2/51)	96% (49/51)	0% (0/26)	100% (26/26)	3% (2/77)	19% (8/43)
MM6/IL-6	100% (30/30)	0% (0/30)	79% (23/29)	17% (6/29)	7% (4/59)	93% (55/59)	0% (0/30)	100% (30/30)	5% (4/89)	10% (6/59)
PBMC/IL-6	100% (30/30)	0% (0/30)	90% (27/30)	10% (3/30)	12% (7/60)	88% (53/60)	0% (0/30)	100% (30/30)	8% (7/90)	5% (3/60)
PBMC/IL-6 (Cryo)[4]	100% (30/30)	0% (0/30)	53% (16/30)	47% (14/30)	10% (6/60)	90% (54/60)	0% (0/30)	100% (30/30)	7% (6/90)	23% (14/60)
WB/IL-6	100% (30/30)	0% (0/30)	93% (27/29)	7% (2/29)	17% (10/59)	83% (49/59)	0% (0/30)	100% (30/30)	11% (10/89)	3% (2/59)
WB/IL-1β (Tube)	97% (28/29)	3% (1/29)	90% (27/30)	10% (3/30)	39% (23/59)	61% (36/59)	3% (1/29)	97% (28/29)	27% (24/88)	7% (4/59)
WB/IL-1β (96-well plate)[5]	100% (28/28)	0% (0/28)	67% (18/27)	33% (9/27)	2% (1/55)	98% (54/55)	0% (0/29)	100% (29/29)	1% (1/84)	16% (9/55)

Abbreviations: Cryo = Cryopreserved; EU/mL = Endotoxin units/mL; IL = Interleukin; MM6 = Mono Mac 6; PBMC = Peripheral blood mononuclear cells; WB = Whole blood

[1]Data shown as a percentage (number of correct, false positive, or false negative runs/total number of runs), based on results of 10 parenteral drugs tested in each of three different laboratories. Samples of each drug were tested with or without being spiked with a Gram-negative endotoxin standard (0, 0.25, 0.5, or 1.0 EU/mL, with 0.5 EU/mL tested in duplicate).

[2]False positive rate = the proportion of all negative substances that are falsely identified as positive.

[3]False negative rate = the proportion of all positive substances that are falsely identified as negative.

[4]A modification of the PBMC/IL-6 test method using cryopreserved PBMCs.

[5]A modification of the WB/IL-1β test method using 96-well plates instead of tubes for the test substance incubation.

[This Page Intentionally Left Blank]

7.0 Reliability of the *In Vitro* Pyrogen Test Methods

An assessment of test method reliability (intralaboratory repeatability and intra- and inter-laboratory reproducibility) is an essential element of any evaluation of the performance of an alternative test method (ICCVAM 2003). Repeatability refers to the closeness of agreement among test results obtained within a single laboratory when the procedure is performed on the same substance under identical conditions within a given time period (ICCVAM 1997, 2003). Intra-laboratory reproducibility refers to the determination of the extent to which qualified personnel within the same laboratory can replicate results using a specific test protocol at different times. Inter-laboratory reproducibility refers to the determination of the extent to which different laboratories can replicate results using the same protocol and test chemicals, and indicates the extent to which a test method can be transferred successfully among laboratories. A reliability assessment includes a quantitative and/or qualitative analysis of intralaboratory repeatability and intra- and inter-laboratory reproducibility. In addition, measures of central tendency and variation are summarized for historical control data (negative, vehicle, positive), where applicable.

An evaluation of intralaboratory repeatability and reproducibility could be conducted because *in vitro* pyrogen test data were available from replicate wells within individual experiments, and from replicate experiments within the individual laboratories. In addition, comparable data were available from each of the three laboratories that performed the validation studies, which allowed an evaluation of interlaboratory reproducibility.

7.1 Selection Rationale for the Substances Used to Evaluate the Reliability of *In Vitro* Pyrogen Test Methods

The quality of a reliability evaluation depends on the extent to which the substances tested adequately represent the range of physicochemical characteristics and response levels that the test method should be capable of evaluating. The rationale for selecting the substances used in the validation studies was discussed in **Section 3.1**. In response to the ICCVAM PWG request for data on other relevant test materials (e.g., medical devices, biologics, etc.) with these test methods, ECVAM summarized published and unpublished studies on snake venom sera, medical devices, dialysate, and lipidic formulations (see question #3 in **Appendix B**).

Each sample contained the appropriate endotoxin spike concentration when tested at its Maximum Valid Dilution (MVD). The MVD takes into account the endotoxin limit concentration (ELC) and the detection limit of the particular test method. The U.S. and European Pharmacopeias assign ELCs for drugs based on their specific administered dose, route of administration, and dosing regimen. Based on the selected threshold pyrogen dose of 0.5 EU/mL (see **Section 4.0**), and the decision criteria used in the validation studies to identify a pyrogenic response (≥0.5 EU/mL, see **Section 5.0**), a concentration of 0.5 EU/mL was used as the detection limit for the *in vitro* test methods when calculating the MVDs for each of the test substances.

7.2 Analysis of Intralaboratory Repeatability and Reproducibility

Intralaboratory repeatability analyses were performed using the OD values obtained for each test with each spiked sample. All analyses of intra- and inter-laboratory reproducibility were performed on the classifications of pyrogenic or non-pyrogenic, rather than on the absolute

OD values generated in each run. Analyses of intra-laboratory reliability include a CV analysis for the log-transformed OD_{450} measurements, which is a statistical measure of the deviation of a variable from its mean (e.g., Holzhütter et al. 1996). According to Section 7.2 of each ECVAM BRD, the analyses focused on the CV because existing data has demonstrated that there is a direct relationship between the mean responses and the variation (e.g., empirical variance or standard deviation). Moreover, the CV should be distributed symmetrically around a constant factor if the mean-variance relationship is linear.

7.2.1 Intralaboratory Repeatability

In the ECVAM validation study, intralaboratory repeatability of each test method was evaluated by testing saline and various endotoxin spikes (0.06 to 0.5 EU/mL) in saline and evaluating the closeness of agreement among OD readings for cytokine measurements at each concentration. Each experiment was conducted up to three times for each test method. Up to 20 replicates per concentration were tested and results indicated that variability in OD measurements increased with increasing endotoxin concentration, but the variability was not so great as to interfere with distinguishing the 0.5 EU/mL spike concentration (i.e., the threshold for pyrogenicity) from the lower concentrations. **Table 7-1** details the study design for each of these evaluations. With the exception of the Cryo WB/IL-1β test method, at least four different study designs were employed for each test method. Appendix C of the ECVAM Cryo WB/IL-1β BRD (see **Appendix A**) indicates that because intralaboratory reliability was extensively evaluated in the WB/IL-1β test method, only a subset (n=2) of these studies was conducted as part of a "catch-up validation" study. Based on the "acceptable" intralaboratory performance in this subset of studies, additional studies were not considered necessary.

With regard to plate-to-plate variation, the ECVAM Trial Data Report (see **Appendix C**) states that the data obtained from each ELISA plate (i.e., 96-well format) must be considered as a whole and cannot be compared to other ELISA plates due to uncontrollable variation. Therefore, it was recommended that each ELISA plate should include all controls (e.g., negative control, positive control, negative product control, and positive product control) required for the analytical procedure.

Table 7-1 Intralaboratory Repeatability Assessed with Saline Spiked with WHO-LPS 94/580

Experiment	Study Design	Test Method				
		MM6/IL-6	PBMC/IL-6	WB/IL-1β	WB/IL-6	Cryo WB/IL-1β[1]
1A	Endotoxin concentration (EU/mL)	0, 0.25, 0.5	0, 0.25, 0.5	0, 0.5	0, 0.5	0, 0.5
	N (per spike)	20	20	32	20	32
	Repetitions of experiment	1	1	1	1	1
1B	Endotoxin concentration (EU/mL)	0, 0.063, 0.125, 0.25, 0.5	0, 0.063, 0.125, 0.25, 0.5	0, 0.063, 0.125, 0.25, 0.5	0, 0.063, 0.125, 0.25, 0.5	0, 0.063, 0.125, 0.25, 0.5
	N (per spike)	12	12	12	10	12
	Repetitions of experiment	1	1	1	1	1
2A	Endotoxin concentration (EU/mL)	0, 0.25, 0.5	0, 0.5	0, 0.5	0, 0.25, 0.5	ND
	N (per spike)	20	8	12	8	ND
	Repetitions of experiment	3	3	3	3	ND
2B	Endotoxin concentration (EU/mL)	0, 0.25, 0.5	0, 0.063, 0.125, 0.25, 0.5	0, 0.25, 0.5	0, 0.5	ND
	N (per spike)	20	8	8	5	ND
	Repetitions of experiment	3	3	3	8	ND
2C	Endotoxin concentration (EU/mL)	ND	0, 0.125, 0.25, 0.5	0, 0.5	ND	ND
	N (per spike)	ND	8	5	ND	ND
	Repetitions of experiment	ND	8	8	ND	ND

Abbreviations: Cryo = Cryopreserved; EU/mL = Endotoxin units/mL; IL = Interleukin; LPS = Lipopolysaccharide; MM6 = Mono Mac 6; N = number of replicates; ND = Not done; PBMC = Peripheral blood mononuclear cells; WB = Whole blood; WHO = World Health Organization

[1]The Cryo WB/IL-1β test method was included in a catch-up validation study to assess intralaboratory reliability in a subset of experiments (n=2).

7.2.2 Intralaboratory Reproducibility

Intralaboratory reproducibility was evaluated using three marketed pharmaceuticals spiked with various concentrations of endotoxin (see **Table 3-2**). Three identical, independent runs were conducted in each of the three testing laboratories, with the exception of the Cryo WB/IL-1β test method.[5] The correlations (expressed as a percentage of agreement) between

[5] The ECVAM Cryo WB/IL-1 test method BRD states that there was no direct assessment of intralaboratory reproducibility because such an evaluation was performed in the WB/IL-1 test method, and the authors assumed that variability would not be affected by the use of cryopreserved blood.

pairs of the independent runs (i.e., run 1 vs. run 2; run 1 vs. run 3; run 2 vs. run 3) were determined and the mean of these three values was calculated. In all reproducibility analyses, a single run consisted of each of the substances assayed in quadruplicate. Acceptability criteria for each run included a CV analysis to remove highly variable responses from the analyses. The criterion used to identify outliers ranged from CV <0.25 to CV <0.45, depending on the method being considered, and was arbitrarily set based on results using saline spiked with endotoxin. As an example, for the MM6/IL-6 test method, the CV for any single spike concentration was ≤ 0.12, and therefore, the outlier criterion was set at 0.25.

Agreement between different runs was determined for each substance in three laboratories. As shown in **Table 7-2**, the agreement across three runs in an individual lab ranged from 75% to 100%

.

Table 7-2 Intralaboratory Reproducibility of *In Vitro* Pyrogen Test Methods

Run Comparison[1]	WB/IL-1β			Cryo WB/IL-1β			WB/IL-6			PBMC/IL-6			MM6/IL-6		
	Lab 1	Lab 2	Lab 3	Lab 1	Lab 2	Lab 3	Lab 1	Lab 2	Lab 3	Lab 1	Lab 2	Lab 3	Lab 1	Lab 2	Lab 3
1 vs 2	92% (11/12)	100% (8/8)	100% (12/12)	ND[3]	ND	ND	75% (9/12)	92% (11/12)	100% (12/12)	92% (11/12)	100% (12/12)	100% (12/12)	100% (12/12)	92% (11/12)	100% (12/12)
1 vs 3	83% (10/12)	88% (7/8)	92% (11/12)	ND	ND	ND	100% (12/12)	92% (11/12)	100% (12/12)	100% (12/12)	100% (12/12)	92% (11/12)	100% (12/12)	92% (11/12)	92% (11/12)
2 vs 3	92% (11/12)	NI[4]	92% (11/12)	ND	ND	ND	75% (9/12)	92% (11/12)	100% (12/12)	92% (11/12)	100% (12/12)	92% (11/12)	100% (12/12)	100% (12/12)	92% (11/12)
Mean	89%	n.c.	95%	ND	ND	ND	83%	92%	100%	95%	100%	95%	100%	95%	95%
Agreement[2] across 3 runs	83%	n.c.	92%	ND	ND	ND	75%	92%	100%	92%	100%	94%	100%	92%	92%

Abbreviations: Cryo = Cryopreserved; IL= Interleukin; MM6 = Mono Mac 6; n.c. = Not calculated; ND = Not done; NI = Not included; PBMC = Peripheral blood mononuclear cells; WB = Whole blood

[1]Comparison among 3 individual runs within each laboratory

[2]All possible combinations of runs among the 3 laboratories were compared.

[3]Not done. The ECVAM Cryo WB/IL-1β BRD states that an assessment of intralaboratory reproducibility was performed using the WB IL-1β (fresh blood) test method, and it was assumed that intralaboratory variability would not be affected by the change to cryopreserved blood assayed in 96-well plates.

[4]Not included due to lack of sufficient data. The sensitivity criteria were not met for 1 of 3 substances in run 2, and 1 of 3 substances in run 3.

[This Page Intentionally Left Blank]

7.2.3 Interlaboratory Reproducibility

Interlaboratory reproducibility was evaluated in two different studies. In both studies, each run from one laboratory was compared with all runs of another laboratory. The proportions of similarly classified samples provide a measure of reproducibility. In the first study, the interlaboratory reproducibility was evaluated using results from three marketed pharmaceuticals spiked with endotoxin and tested in triplicate in each of the three laboratories. As shown in **Table 7-3**, the agreement across three laboratories for each test method (where three runs per laboratory were conducted) ranged from 58% to 86%, depending on the test method considered. In comparison, the agreement across three laboratories for the Cryo WB/IL-1β test method, for which only one run per laboratory was conducted, was 92%.

Table 7-3 Interlaboratory Reproducibility of *In Vitro* Pyrogen Test Methods

Lab Comparison[1]	Agreement Between Laboratories[1]				
	WB/IL-1β (Tube)	Cryo WB/IL-1β	WB/IL-6	PBMC/IL-6	MM6/IL-6
1 vs 2	92% (77/84)[2]	92% (11/12)[3]	72% (78/108)	81% (87/108)	97% (105/108)
1 vs 3	77% (83/108)	92% (11/12)[3]	75% (81/108)	86% (93/108)	89% (96/108)
2 vs 3	68% (57/84)[2]	92% (11/12)[3]	97% (105/108)	89% (96/108)	86% (93/108)
Mean	79%	92%	81%	85%	90%
Agreement across 3 labs[4]	58% (167/288)[2]	92% (11/12)[3]	72% (234/324)	78% (252/324)	86% (279/324)

Abbreviations: Cryo = Cryopreserved; IL= Interleukin; MM6 = Mono Mac 6; PBMC = Peripheral blood mononuclear cells; WB = Whole blood

[1]Data from three substances (see **Table 3-2**) spiked with endotoxin (WHO-LPS 94/580 [*E. coli* O113:H10:K-]) at 0, 0.5, and 1.0 endotoin units/mL (EU/mL), with 0 EU/mL tested in duplicate, were tested three times in three different laboratories, with the exception of Cryo WB/IL-1β (only the preliminary run from each laboratory used for analysis).

[2]Some of the runs did not meet the assay acceptance criteria and therefore were excluded from the analysis.

[3]For the Cryo WB/IL-1β test method, each substance tested only once in each laboratory.

[4]All possible combinations of runs among the 3 laboratories were compared (with the exception of Cryo WB/IL-1β, which was only tested once in each laboratory, resulting in only one possible combination per substance).

In the second study, interlaboratory reproducibility was evaluated with the same 10 substances used for evaluating accuracy. In this study, each of the substances was spiked with four concentrations of endotoxin (with one conentration spiked in replicate) and tested once in each of three laboratories. As shown in **Table 7-4**, the agreement across three laboratories for each test method ranged from 57% to 88%, depending on the test method considered. The extent and order of agreement among laboratories was the same for both studies; the WB/IL-1β tube method showed the least agreement (57-58%) and the Cryo WB/IL-1β test method showed the most (88-92%).

Table 7-4 Interlaboratory Reproducibility of *In Vitro* Pyrogen Test Methods

Lab Comparison[1]	Agreement Between Laboratories[1]						
	WB/IL-1β (Tube)	WB/IL-1β (Plate)	Cryo WB/IL-1β	WB/IL-6	PBMC/IL-6	PBMC/IL-6 (Cryo)	MM6/IL-6
1 vs 2	73% (35/48)	88% (37/42)	84% (38/45)	85% (41/48)	84% (42/50)	96% (48/50)	90% (45/50)
1 vs 3	82% (40/49)	90% (35/39)	88% (21/24)	85% (41/48)	86% (43/50)	76% (38/50)	90% (43/48)
2 vs 3	70% (33/47)	92% (43/47)	100% (25/25)	88% (44/50)	90% (45/50)	80% (40/50)	83% (40/48)
Mean	75%	90%	91%	86%	87%	84%	88%
Agreement across 3 labs	57% (27/47)	85% (33/39)	88% (21/24)	79% (38/48)	80% (40/50)	76% (38/50)	81% (39/48)

Abbreviations: Cryo = Cryopreserved; IL = Interleukin; MM6 = Mono Mac 6; PBMC = Peripheral blood mononuclear cells; WB = Whole blood

[1]Data from 10 substances spiked with endotoxin (WHO-LPS 94/580 [*E. coli* O113:H10:K-]) at 0, 0.25, 0.5, and 1.0 endotoxin units/mL (EU/mL), with 0.5 EU/mL tested in duplicate, were tested once in three different laboratories.

7.3 Historical Positive and Negative Control Data

No historical control data were provided for any of the five *in vitro* pyrogen test methods. However, the intralaboratory repeatability analysis described in **Section 7.2.1** included repeat testing of both spiked (0.5 EU/mL endotoxin) and non-spiked saline, and the accumulated positive and negative control values, respectively for each of the methods. As a result, the database that was accumulated during the ECVAM validation studies provides an indication of the range and variability in responses for the positive and negative controls.

8.0 Test Method Data Quality

8.1 Adherence to National and International GLP Guidelines

Ideally, all data supporting the validity of a test method should be obtained and reported in accordance with GLP guidelines (i.e., OECD 1998; EPA 2003a, 2003b; FDA 2003). These guidelines provide an internationally standardized approach for the reporting requirements of studies designed for regulatory submissions, internal audits of laboratory records and data summaries, the archive of study data and records, and information about the test protocol and laboratory personnel, to provide assurances regarding the integrity, reliability, and accountability of the study.

The initial ECVAM validation studies for the five *in vitro* pyrogen test methods were conducted "in the spirit of" GLP requirements (i.e., written protocols and approved SOPs were followed during the entire course of the study). In the catch-up validation studies, two GLP laboratories and two National Control Laboratories participated.

8.2 Data Quality Audits

Formal assessments of data quality, such as a QA audit, generally involve a systematic and critical comparison of the data provided in a study report with the laboratory records generated for the study. No attempt was made to formally audit the quality of the data presented in the five ECVAM BRDs. However, as indicated in **Section 5.2**, the raw data from the validation studies are available from the participating laboratories for a quality analysis.

8.3 Impact of Deviations from GLP Guidelines

The impact of the deviations from the GLP guidelines, as reported in the ECVAM BRDs, was not evaluated.

8.4 Availability of Laboratory Notebooks or Other Records

All records are stored and archived by the participating laboratories and are available for inspection.

8.5 Need for Data Quality

Data quality is a critical component of the validation process. To ensure data quality, ICCVAM recommends that all data generated during the validation of a test method be available, along with the detailed protocol(s) under which the data were produced. Original data should be available for examination, as should supporting documentation such as laboratory notebooks. Ideally, the data should adhere to GLP guidelines (ICCVAM 1997). Data protocols for the validation studies summarized here are available from ECVAM (see **Appendix A**), and the data from the individual laboratories are available for inspection, as indicated in **Section 8.4**.

[This Page Intentionally Left Blank]

9.0 Other Scientific Reports and Reviews

The individual BRDs submitted by ECVAM (i.e., one for each of the *in vitro* pyrogen test methods) are provided in **Appendix A** and were used in the performance analyses described in **Section 6.0** and **Section 7.0**. A *FR* notice (Vol. 70, No. 241, pp. 74833-74834, December 16, 2005) was published requesting the submission of data from the RPT, the BET, or an *in vitro* pyrogen test method. No data were received in response to this request.

NICEATM conducted a prescreen evaluation of the ECVAM BRDs to verify that the information contained within the documents fulfilled the requirements outlined in the ICCVAM submission guidelines (ICCVAM 2003). Based on this evaluation, the ICCVAM PWG requested a direct comparison on the accuracy analysis of the *in vitro* test methods with the reference test methods (i.e., the RPT and the BET) and data to support the claim that the *in vitro* test methods can detect non-endotoxin pyrogens. In response to these requests, ECVAM provided supplemental data from published (e.g., Francois et al. 2006) and unpublished studies in an attempt to address these issues (see questions #1 and #2 in **Appendix B**).

Hartung et al. (2001) provided a summary report of an ECVAM-sponsored workshop to review the current status of pyrogenicity testing, to review the capabilities of new pyrogen tests, and to provide recommendations for their continued development. The need for alternatives to the RPT and the BET was discussed, and their respective limitations were highlighted. The workshop compared the utility of the various methods (i.e., *in vitro* pyrogen test methods, BET, RPT) for testing a variety of pyrogenic materials. Workshop conclusions indicated a need for alterative test methods to address the limitations of the BET and RPT, but stressed the need for appropriate validation of any new method.

9.1 Summaries of *In Vitro* Pyrogen Test Methods and Data from Published and Unpublished Studies

As indicated in **Section 1.5**, NICEATM conducted an online literature search for relevant information on the proposed test methods using multiple databases (i.e., PubMed, SCOPUS, TOXLINE, Web of Science). This search revealed ten additional scientific publications that contained data from *in vitro* pyrogen product testing. These studies contained comparisons of the results obtained in an *in vitro* test method with those obtained in the RPT and/or BET (see **Tables 9-1** to **9-8**). These studies were not included in previous sections of the ICCVAM BRD because they used a different method or protocol, or because they lacked sufficient information for an evaluation of accuracy and reliability (e.g., an adequate validation study design was not included, a standardized reference pyrogen was not used). Summaries of these published studies and available data from the *in vitro* pyrogen methods are presented below.

9.1.1 Andrade et al. (2003)

The authors evaluated the utility of human PBMCs and diluted WB for *in vitro* pyrogen tests and compared the responses to those obtained in the BET and RPT for the same diverse sampling of parenteral pharmaceuticals and biological products (see **Tables 9-1** and **9-2**). Interference testing of each substance was performed with spikes of the international endotoxin standard WHO-LPS 94/580. These studies established an endotoxin detection limit

of 0.06 EU/mL for both *in vitro* assays, and the results were consistent with those from the BET and RPT. The authors concluded that both the PBMC and WB methods were comparable to the BET and the RPT in their ability to detect and quantify the presence of endotoxin. In addition, the WB test method was able to detect concentration-dependent IL-6 release on exposure of WB to non-endotoxin pyrogens and pyrogens from Gram-positive organisms (i.e., *Candida albicans* and *Staphylococcus aureus*).

Table 9-1 Results of Pyrogen Testing of Pharmaceutical/Biological Products in the Human PBMC Assay, the BET, and the RPT[1]

Product	Number of Batches[2]	PBMC (EU/mL)	BET (EU/mL)	RPT
Ampicillin - 1000 mg/5 mL A	1	<6	<0.06	Pass
Ampicillin - 1000 mg/5 mL A	1	<6	<0.06	Pass
Gentamycin - 80 mg/2 mL	2	<3	<0.06	Pass
Oxacillin - 500 mg/5 mL	2	<3	<0.06	Pass
Enoxaparin - 100 mg/mL	3	<1.2	<0.06	Pass
Insulin - 100 U/mL	2	<3	<0.06	Pass
Tenoxican - 40mg/2 mL	1	<6	<0.06	Pass
Metoclopramide - 10 mg/2 mL	4	<3	<0.06	Pass
Calcium folinate - 50 mg/5 mL	1	<2.4	<0.06	Pass
Ranitidine - 25 mg/mL	2	<6	1.2-2.4	Pass
Pantoprazol - 40 mg/10 mL	1	<3	<0.06	Pass
Human serum albumin - 20%	1	<4.8	0.48-0.96	Pass
Erythropoietin - 4000 IU/vial A	1	<1.2	0.48-0.96	Pass
Erythropoietin - 2000 IU/vial B	1	112 ± 10[1]	491-983	Fail
Erythropoietin - 4000 IU/vial C	1	<1.2	<0.06	Pass
recG-CSF - 200 μg/vial A	3	<0.6	<0.06	Pass
Saline solution - 0.9% A	1	<0.3	<0.06	Pass

Abbreviations: BET = Bacterial Endotoxin Test; CSF = Colony Stimulating Factor; EU/mL = Endotoxin units/mL; IU = international units; PBMC = Peripheral blood mononuclear cells; rec = Recombinant; RPT = Rabbit pyrogen test; U = units

[1]From Andrade et al. (2003)

[2]Batch results were combined; PBMC and BET study values represent a mean ± standard deviation value or consensus detection limits (n=3 donors; 4 replicates from each donor).

Table 9-2 Results of Pyrogen Testing of Pharmaceutical/Biological Products by the Human WB Culture Assay, the BET, and the RPT[1]

Product	Number of Batches[2]	WB Culture (EU/mL)	BET (EU/mL)	RPT
Dipyrone - 500 mg/mL	3	<24	<0.06	Pass
Amikacin - 500 mg/2 mL	2	<12	<0.06	Pass
Ampicillin - 1000 mg/5 mL A	1	<6	<0.06	Pass
Ampicillin - 1000 mg/5 mL A	1	<6	<0.06	Pass
Gentamycin - 80 mg/2 mL	2	<6	<0.06	Pass
Oxacillin - 500 mg/5 mL	2	<6	<0.06	Pass
Vancomycin - 500 mg/5 mL	2	<6	<0.06	Pass
Enoxaparin - 100 mg/mL	3	<0.6	<0.06	Pass
Heparin - 5000 IU/mL	2	<0.6	<0.06	Pass
Insulin - 100 U/mL	3	<6	<0.06	Pass
Ketoprofen - 100 mg/2mL	1	<6	<0.06	Pass
Diclofenac - 75 mg/3 mL	1	<12	<0.06	Pass
Tenoxicam - 40 mg/2 mL	2	<6	<0.06	Pass
Metoclopramide - 10 mg/2 mL	3	<3	<0.06	Pass
Cytarabine - 100 mg/5mL	1	<1.2	<0.06	Pass
Calcium folinate - 50 mg/5 mL	1	<0.6	<0.06	Pass
Ranitidine - 25 mg/mL	1	<6	1.2-2.4	Pass
Pantoprazol - 40 mg/10 mL	1	<6	<0.06	Pass
Furosemide - 10 mg/mL	2	<0.6	<0.06	Pass
rec-hGH - 4 IU/vial A	2	<0.2	<0.06	Pass
rec-hGH - 4 IU/vial B	1	12.4±2.5[1]	15.84-31.68	Pass
Human serum albumin - 20%	1	<2.4	0.48-0.96	Pass
Erythropoietin - 4000 IU/vial A	1	0.76	0.48-0.96	Pass
Erythropoietin - 2000 IU/vial B	1	141±2.8[1]	491-983	Fail
Erythropoietin - 4000 IU/vial C	1	<0.6	<0.06	Pass
recG-CSF - 300 μg/vial	3	<0.6	<0.06	Pass
Saline solution 0.9% A	2	<0.3	<0.06	Pass
Saline solution 0.9% B	1	44.8±5[1]	48-96	Fail
Glucose - 0.5%	1	2054±95[1]	1920-3840	Fail
Vitamin K - 10 mg/mL	2	<6	<0.06	Pass

Abbreviations: BET = Bacterial endotoxin test; CSF = Colony stimulating factor; EU/mL = Endotoxin units/mL; hGH = Human growth hormone; IU = International units; rec = Recombinant; RPT = Rabbit pyrogen test; WB = Whole blood; U = units

[1]From Andrade et al. (2003)

[2]Batch results were combined; PBMC and BET study values represent a mean ± standard deviation value or consensus detection limits (n= 3 donors; 4 replicates from each donor).

9.1.2 Bleeker et al. (1994)

This study measured IL-6 release from PBMCs as an indicator of pyrogenicity for *in vitro* safety testing of hemoglobin (Hb) solutions. The authors demonstrated that pure, polymerized Hb produced under aseptic conditions did not induce or inhibit IL-6 production, whereas production under non-aseptic conditions led to IL-6 release, which was also seen with the BET. Based on these findings, the authors concluded that IL-6 release from isolated PBMCs provides a sensitive indicator of endotoxin contamination in Hb solutions. The observed detection limit for endotoxin in Hb solutions (below 0.4 EU/mL) led the authors to suggest that this test method would be more sensitive to the presence of endotoxin than the RPT.

9.1.3 Carlin and Viitanen (2003)

Using WB and MM6-based *in vitro* pyrogen methods, this study evaluated the pyrogenic potential of a multivalent vaccine, Infanrix® (GlaxoSmithKline) that contains protein and polysaccharide components from both Gram-positive and Gram-negative bacteria. The five Infanrix® vaccines studied (e.g., Infanrix®, Infanrix® Hep B, Infanrix® polio, Infanrix® hexa, and Infanrix® polio Hib) contain Gram-positive bacterial components that are potentially pyrogenic but not detectable in the BET. IL-6 production in the WB/IL-6 test method varied among the seven donor blood samples in response to each of the five vaccines. Some donor samples produced a weak or no IL-6 release and others produced a large release (**Table 9-3**). However, IL-6 production from any single donor was similar for all vaccines when tested at various times. The variability in the magnitude of response to each vaccine among donors and the consistency of the response of any single donor was also seen when IL-1β was used as a marker. IL-6 release from WB was also examined following exposure to three concentrations of endotoxin standard (0.2, 2, and 20 pg/1.2 mL). All donor WB samples released IL-6 in a concentration responsive manner.

The IL-6 release from MM6 cells (**Table 9-4**) exposed to the five Infanrix® vaccines was measured using an ELISA and compared to the responses induced by three concentrations of endotoxin standard (0.2, 2, and 20 pg/1.2 mL) in three separate experiments. The MM6 cells produced minimal responses to the vaccines when compared to WB, but released significant amounts of IL-6 in response to high concentrations of endotoxin. However, IL-6 induction by two different endotoxin standards in MM6 cells was strongly attenuated (>80% inhibition) when either of two vaccines (Infanrix® and Infanrix® Hep-B) was present (data not included in **Table 9-4**). Based on these studies, the authors suggested that a BET or RPT result might not correlate with the human fever response one might expect in humans immunized with such vaccines, because the production of proinflammatory cytokines may be compromised by various components in the vaccine product, and because Gram-positive components in the vaccines would not be detected in the BET.

Table 9-3 IL-6 Production from WB after Exposure to Endotoxin or Five Infanrix® Vaccines[1,2]

Experiment (Blood Donor)	Endotoxin		Vaccine (Absorbance in ELISA; n=4[3])					
	Endotoxin (pg/1.2 mL)	Absorbance (ELISA)	Dilution (µL vaccine/1.2 mL	Infanrix®	Infanrix® Hep-B	Infanrix® Hexa	Infanrix® Polio	Infanrix® Polio Hib
1	0.2	0.47	0.03	0.945	1.052	1.069	0.869	1.082
	2	0.971	0.3	1.826	2.055	2.014	1.832	1.919
	20	1.116	3	2.826	2.587	2.638	2.609	2.2
2	0.2	0.001	0.03	0.149	0.256	0.231	NT	0.284
	2	0.127	0.3	0.869	0.847	1.095	NT	0.933
	20	0.764	3	1.998	1.986	2.187	NT	1.685
3	0.2	-0.007	0.03	0.005	0.037	0.009	0.007	0.208
	2	0.09	0.3	0.275	0.457	0.282	0.321	0.261
	20	0.811	3	0.941	1.057	0.795	1.284	1.325
4	0.2	0.006	0.03	0.056	0.053	0.028	0.088	0.104
	2	0.043	0.3	0.165	0.312	0.44	0.309	0.533
	20	0.458	3	1.229	1.489	1.476	1.181	1.242
5	0.2	0.043	0.03	-	0.071	-0.003	-0.003	0.011
	2	0.024	0.3	0.007	0.014	0.004	0.03	0.05
	20	0.435	3	0.042	0.164	0.008	0.08	0.12
6	0.2	0.013	0.03	-0.009	-0.018	-0.01	-0.022	0.012
	2	0.022	0.3	-0.007	-0.008	0.005	-0.019	-0.007
	20	0.569	3	0.132	0.411	0.042	0.132	0.188
7	0.2	0.036	0.03	-0.012	-0.012	-0.01	-0.014	0.07
	2	0.014	0.3	-0.01	-0.01	-0.012	-0.011	-0.013
	20	0.436	3	0.183	0.274	0.045	0.183	0.525

Abbreviations: ELISA = Enzyme-linked immunosorbent assay; Hep = Hepatitis; IL-6 = Interleukin-6; NT = Not tested; WB = Whole blood

[1]From Carlin and Viitanen (2003)

[2]WB was challenged with endotoxin standard or vaccine in pyrogen-free water to provide the final concentration and incubated overnight at 37°C.

[3]Duplicate samples were run in two separate experiments.

Table 9-4 IL-6 Production by MM6 Cells after Exposure to Endotoxin or Five Infanrix® Vaccines[1,2]

MM6 Batch	Endotoxin			Vaccine (Absorbance in ELISA; 250,000 MM6 cells); n=4[3]				
	Endotoxin (pg/1.2 mL)	Absorbance in IL-6 ELISA	Dilution (µL vaccine/ 1.2 mL	Infanrix®	Infanrix® Hep-B	Infanrix® Hexa	Infanrix® Polio	Infanrix® Polio Hib
1	0.2	-0.001	0.3	0.013	0.014	0.001	0.002	-0.001
	2	0.026	3	0.078	0.158	0.06	0.105	0.07
	20	0.383	30	0.054	0.052	0.053	0.106	0.089
2	0.2	-0.001	0.3	0.004	0.01	0.001	0.003	0.004
	2	0.025	3	0.033	0.062	0.019	0.037	0.032
	20	0.4	30	0.013	0.012	0.018	0.038	0.038
3	0.2	-0.009	0.3	-0.012	-0.017	-0.021	-0.014	-0.019
	2	0.03	3	0.019	0.05	0.01	0.043	0.026
	20	0.192	30	-0.018	-0.012	-0.007	0	0.005

Abbreviations: ELISA = Enzyme-linked immunosorbent assay; IL-6 = Interleukin-6; MM6 = Mono Mac 6

[1]From Carlin and Viitanen (2003)

[2]MM6 cells were stimulated with endotoxin standard or vaccine in pyrogen-free water to provide the final concentration and incubated overnight at 37°C.

[3]n = Duplicate samples were run in two separate experiments.

9.1.4 Carlin and Viitanen (2005)

This study provides support for the findings from a previous study (Carlin and Viitanen (2003) in which the authors demonstrated IL-6 release by a WB method in response to pyrogenic or spiked multivalent vaccine preparations that were inactive in the BET. It also confirms that IL-6 was released from WB of some, but not all donors. The present study demonstrates that IL-6 release in susceptible donors was caused by toxoids from Gram-positive diphtheria, and to a lesser extent, from tetanus bacterial components of the vaccines. The WB donors were studied for two years and their responses to the individual vaccines, whether responsive or non-responsive, were consistent. The responses of these donors to Gram-negative endotoxin or lipoteichoic acid (LTA) from Gram-positive bacteria were consistent and confirmed the findings of Fennrich et al. (1999) with respect to the consistency of responses among several hundred blood donors to endotoxin. The authors concluded that individual donor-specific differences in IL-6 release from WB exposed to the multivalent vaccines resulted from toxoids present in the diphtheria or tetanus component, and noted that these donor-specific responses to the vaccines were not observed in the BET.

9.1.5 Daneshian et al. (2006)

This study describes the development of a modification to the WB/IL-1β method termed AWIPT (Adsorb, Wash, *In Vitro* Pyrogen Test). The authors indicate that this modification is intended to increase sensitivity to the presence of endotoxin contamination by isolating endotoxin from WB. To accomplish this, the sample containing endotoxin (naturally occurring or spiked) is treated with human serum albumin (HSA) covalently linked to macroporous acrylic beads. The HSA-treated beads bind the endotoxin, which is subsequently eluted from the beads. The WB/IL-1β test method is therefore performed using a slightly modified protocol in which the diluted WB is incubated overnight with the sample in the bead suspension.

The results showed that HSA-coated beads bind endotoxin in a concentration-dependent manner (when spiked with 0, 25, 50, and 100 pg/mL LPS), but little or none was bound to unmodified beads. The test showed a detection limit of 25 pg/mL LPS (i.e., 0.25 EU/mL), which is less sensitive than the BET (3 pg/mL) and more sensitive than the RPT (50 pg/mL). IL-1β secretion in response to either LPS or LTA was generally higher using the AWIPT procedure, but the concentrations of LPS or LTA needed to induce a response were similar; thus the sensitivity of this test modification was comparable to that of the unmodified WB/IL-1β test method.

Daneshian et al. studied the kinetics of cytokine release from WB in response to a challenge with 2 pg/mL of endotoxin. IL-1β release in the AWIPT-treated samples lagged slightly behind that of the standard WB/IL-1β test in the 0 to 8 hr time period, whereas more IL-1β was produced in the AWIPT-treated samples in the 10 to 30 hr time period. Some immunomodulatory or toxic cancer drug samples tested in the WB/IL-1β method interfered with the WB/IL-1β assay and required a higher dilution (1/10 to 1/100) to detect IL-1β. Detection of endotoxin spiked into these test samples (measured as IL-1β release) generally occurred at lower dilutions in AWIPT than in the WB/IL-1β test method, suggesting that the interfering substances were removed by the procedure. For example, five dilutions (ranging from 1/3 to 1/316) of liposomal daunorubicin were spiked with 25 pg/mL of endotoxin and

detection of IL-1β was compared between the two methods. This cytokine was not detectable in the WB/IL-1β method (< 30% of the IL-1β released by endotoxin) at any drug dilution, whereas in the AWIPT, IL-1β was detected at drug dilutions of 1/32, 1/100, and 1/316 (>78% of the IL-1β released by endotoxin).

The authors concluded that the inclusion of endotoxin adsorption and washing steps in the WB/IL-1β method (i.e., the AWIPT) to remove potentially interfering substances improved the detection of pyrogenic contaminants in immunomodulatory and toxic cancer drug samples. They suggest that the AWIPT method offers an improvement for safety testing of products administered to patients, and for batch control in pharmaceutical processing.

9.1.6 Eperon et al. (1996, 1997)

Eperon and colleagues developed an *in vitro* test system for measuring pyrogenic substances using two clones derived from MM6 cells (Professor Ziegler-Heitbrock, University of Munich) and one from a THP-1 cell line (European Collection of Animal Cell Cultures, Porton Down, Salisbury, U.K.). These clones are reported to be more phenotypically stable over time with respect to their superior responsiveness to endotoxin than the parent cell lines. Endotoxin content was measured by the release of TNF-α using an immunoassay. These clones demonstrate high LPS sensitivity when non-pyrogenic fetal calf serum is used in the assay as a serum supplement. Enhanced expression of the cell-surface endotoxin receptor CD14 was obtained by pretreatment of the cells for two days with calcitrol. Purified endotoxin (i.e., LPS; smooth strain and rough mutant), other cellular components from Gram-negative or Gram-positive bacteria, and Mycobacteria were tested. The MM6 clones responded to these pyrogenic products in an order of potency of detection equivalent to that found in the RPT and similar to that observed in the BET (i.e., Gram-negative endotoxin > Gram-positive material > non-endotoxin pyrogens). The response of the THP-1 clone was similar to that of the MM6 clones, except that the THP-1 clone did not respond to diphosphoryl lipid A, a structural component of LPS.

Pyrogen testing of a panel of stable blood products, including albumin and Immunoglobulin G (IgG) for parenteral use, produced similar results in the RPT and MM6 or THP-1 clones when tested as received (i.e., free of detectable pyrogens). The products produced positive results when spiked with 20 EU/mL of endotoxin (**Table 9-5**), with a few exceptions. For example, in the cell-based test, there was one borderline but significantly positive result in an unspiked sample, representing a false positive result relative to the RPT. In the BET, 4 of 13 (31%) unspiked samples tested positive (i.e., false positive). The results suggest that the cell-based assays may produce fewer false positives than the BET.

When 10 bacterial and viral vaccine preparations were evaluated, the monocytoid cell-based test method (e.g., combined results from two experiments with each cell line) correlated well with the RPT (positive or negative for endotoxin) with the exception of one preparation that produced nearly 10-fold less TNF-α than the other samples, and was near the limit of detection. This result was not significantly different from the negative control (**Table 9-6**). The authors suggest that these cloned monocytoid cell-based test methods are valid *in vitro* alternatives for detection of endotoxin in commercial preparations, and produce results comparable to the RPT and BET.

Table 9-5 Pyrogenic Activity of Blood Preparations for Parenteral Use[1]

Preparation	Endotoxin Spike	RPT[2]	BET[3,4]	Cell Test[5,6]
IgG for i.v. use	20 EU/mL	+	+	+
	20 EU/mL	+	+	+
	-	-	-	-
	-	-	-	-
	-	-	-	-
	-	-	-	-
	-	-	-	-
	-	-	-	-
	-	-	-	-
	-	-	-	-
Albumin	20 EU/mL	+	+	+
	20 EU/mL	+	+	+
	-	-	-	-
	-	-	-	-
	-	-	-	-
	-	-	-	-
	-	-	-	-
Test threshold[7]		*Δt = 1.5 °C*	*300 pg/mL LPS*	*50 pg/mL TNF*

Abbreviations: BET = Bacterial endotoxin test; EU/mL = Endotoxin units/mL; IgG = Immunoglobulin G; i.v. = Intravenous; LPS = Lipopolysaccharide; RPT = Rabbit pyrogen test; TNF = Tumor Necrosis Factor

[1]From Eperon et al. (1997)

[2]n=3

[3]n=2

[4]Haemachem BET (St. Louis)

[5]n=4 [Note: Cell type not specified; author claims that the Mono Mac 6 or acute monocyte leukemia THP-1 cell lines are equally capable of endotoxin detection.]

[6]TNF induction was determined using a commercial TNF Enyme-linked immunosorbent assay.

[7]RPT threshold was obtained from the European Pharmacopeia; the threshold for the BET and cell-based test methods was considered to be equal to 2 standard deviations from the mean of a set of negative samples.

Table 9-6 Pyrogenic Activity of Vaccine Preparations[1]

Vaccine Preparation[2]	Batch[3]	Cell Test[4]	Pyrogenicity
IgG for i.m. use	A-1	n.d.[5]	-
	A-2	n.d.	-
Bacterial vaccines	B-1	10.8±0.3	+
	C-1	6.0±3.6	+
	D-1	1.4±1.8	-
Viral vaccines	E-1	n.d.	-
	E-2	n.d.	-
	F-1	n.d.	-
	F-2	n.d.	-
	G-1	21.2±3.2	+

Abbreviations: IgG = Immunoglobulin G; i.m. = Intramuscular; n.d. = Non-detectable
[1]From Eperon et al. (1997)
[2]Vaccine solutions were tested at 1/20 (v/v)
[3]Letters refer to distinct types of vaccine preparations; numbers to different lots
[4]Tumor necrosis factor (TNF)-α production in ng/mL±standard error of the mean (n=3) [Note: Cell type not specified; author claims that the Mono Mac 6 or acute monocyte leukemia THP-1 cell lines are equally capable of endotoxin detection.]
[5]No measurable quantity of cytokine was detected.

9.1.7 Marth and Kleinhappl (2002)

In 2000, Ticovac®, a thiomersal- and albumin-free tick-borne encephalitis (TBE) vaccine, was developed as a more immunogenic alternative to previous vaccines that also produced fewer side effects. Although the Austrian health authorities approved this vaccine, 779 cases of fever were reported in children less than 15 years of age, including a high incidence of febrile convulsions in children ages 2 and younger. To determine the cause of these fever reactions, Ticovac® was compared to FSME-Immun®, a TBE-vaccine that rarely resulted in febrile reactions, in an *in vitro* human WB assay that measured cytokine release (i.e., IL-1β, IL-6, IL-8, and TNF-α) as an indication of immune system activation. Ticovac®, which differs from FSME-Immun® only in the albumin component, induced high amounts of TNF-α ($P \leq 0.0001$) and lower amounts of IL-1β ($P \leq 0.05$) as compared to FSME-Immun®. The addition of 0.5 mg of albumin (i.e., the identical quantity of albumin in FSME-Immun®) to Ticovac® reduced the TNF-α induction significantly, resulting in TNF-α production that was similar to the level stimulated by FSME-Immun®. The incubation of Ticovac® with human WB resulted in an increase in TNF-α concentration after 4 hr (peaking at 15 hr) and returned to baseline levels by 27 hr. IL-1β release displayed a similar time course. This temporal response to Ticovac® correlated well with the progression of the clinical outcome (i.e., fever and convulsions in children 6 to 8 hr after the first immunization). Although the mechanism of cytokine production by Ticovac® is unknown, it is clearly linked to the absence of albumin, which is needed as a stabilizer to bind to the antigen of the vaccine. Thus, it was recommended that albumin be added to subsequently produced TBE-vaccines to inhibit nonspecific, excessive immunological reactions.

9.1.8 Martis et al. (2005)

The goal of this study was to establish the cause of 186 cases of aseptic peritonitis that occurred between 2001 and 2003 in peritoneal dialysis patients using an icodextrin-containing dialysate that met both European and USP standards. These patients were not febrile or toxic in appearance, but abdominal pain that was modest to absent and cloudy dialysate were common features. The authors conducted physical, chemical, and microbiological analyses on the recalled dialysate and calculated dose-response curves for IL-6 production in PBMCs from human donors and for sterile peritonitis in rats. Increased levels of IL-6 were identified in dialysis solutions of compliant batches (n=3), but not in non-complaint batches (n=2). Effluents from compliant batches also stimulated IL-6 release in the PBMC assay. Polymyxin B did not inhibit this response, suggesting that a lipopolysaccharide was not responsible for the increased IL-6 levels. When neither Gram-negative nor Gram-positive bacterial contamination was identified in the dialysates, the possibility of a non-endotoxin contaminant was considered. A Gram-positive bacterial cell wall component (i.e., peptidoglycan [PG]) was identified using a silkworm larvae assay in a significant number of dialysates. In the PBMC assay, IL-6 release increased with PG concentration in a dose-response manner. A microbial investigation revealed that the dialysates were contaminated with a Gram-positive organism (i.e., *Alicyclobacillus acidocaldarius*), which contains approximately 40% PG in its cell wall. In rat studies, intraperitoneal injection of icodextrin containing PG (0 - 5000 μg/mL) produced a dose-dependent inflammatory response as measured by an increase in TNF-α and IL-6 production. Subsequent PG contamination has been eliminated using more stringent filtration and carbon treatment steps in the manufacturing process, assaying for PG contamination with the silkworm larvae test, and measuring IL-6 production with the PBMC assay. The lack of aseptic peritonitis incidents that have occurred since the implementation of these additional detection processes support the concept that PG contamination of dialysate was responsible for the reported cases of aseptic peritonitis.

9.1.9 Pool et al. (1998)

This study describes a WB assay for the detection of pyrogens in blood products. IL-6 release from WB in response to endotoxin is used to define a pyrogenic response. This assay was highly responsive to *E. coli* endotoxin (i.e., the limit of detection of endotoxin was 1.25 EU/mL), and also responded to whole bacteria (*E. coli* and *Bacillus subtilis*). There was considerable variation in IL-6 levels released from WB between donors following exposure to endotoxin, but each donor response was always linear. The potential pyrogenicity of production batches of HSA, fibronectin (Fn), and stabilized human serum (SHS) solutions were evaluated using the WB method and compared to the BET and RPT. Spike recovery in batches of these samples varied between 90 and 116% for *E. coli* endotoxin, 74 to 111% for *B. subtilis*, and 61 to 99% for *E. coli* and the products tested did not interfere with the IL-6 assay system. Good correlations were found among the WB, BET, and RPT results (**Table 9-7**). Of 22 products tested, the WB assay and the RPT were in agreement (i.e., pass or fail) for all tests, while one sample was classified as negative in the BET, but positive in both the WB method and the RPT. The detection limit for endotoxin by the WB method was 1.25 EU/mL, which is lower than the established pyrogen cut-off level (as stated in the European Pharmacopeia) for the products under investigation (i.e., 2 EU/mL for HSA and SHS; 4.5

EU/mL for Fn). The authors concluded that the WB assay was able to detect both Gram-negative and a Gram-positive pyrogens and exhibited greater sensitivity to endotoxin than the RPT.

Table 9-7 Comparison of the WB test, BET, and the RPT for Detecting Pyrogens in Production Batches of Biological Products[1]

Product	Batch	WB (EU/mL)[2]	BET	RPT
Fibronectin - 0.5 mg/mL	Fn3195	<0.05	Pass	Pass
	Fn3296	<0.05	Pass	Pass
	Fn3596	1.28	Pass	Pass
Human serum albumin - 200 mg/mL	B274	29.4	Fail	Fail
	B291	<0.05	Pass	Pass
	B293	<0.05	Pass	Pass
	B294	<0.05	Pass	Pass
	B295	<0.05	Pass	Pass
	B296	<0.05	Pass	Pass
	B297	<0.05	Pass	Pass
	B298	1	Pass	Pass
	B299	1.1	Pass	Pass
	B300S	1	Pass	Pass
	B301	<0.05	Pass	Pass
	B302	>20	Pass[3]	Fail
Stabilized human serum - 50 mg/mL	SS349	0.7	Pass	Pass
	SS350	<0.05	Pass	Pass
	SS351	<0.05	Pass	Pass
	SS352	0.5	Pass	Pass
	SS353	<0.05	Pass	Pass
	SS354	0.6	Pass	Pass
	SS355	0.5	Pass	Pass

Abbreviations: BET = Bacterial endotoxin test; EU/mL = Endotoxin units/mL; RPT = Rabbit pyrogen test; WB = Whole blood

[1]From Pool et al. (1998)

[2]Result based on interleukin-6 secretion in human WB using an enzyme-linked immunosorbent assay calibrated to an *E. coli* endotoxin standard (Kabi Diagnostica).

[3]False negative relative to the RPT response

9.1.10 Taktak et al. (1991)

This paper summarizes the development of an *in vitro* pyrogen test method based on IL-6 release from MM6 cells. A detectable level of IL-6 was released in response to 2.5 pg/mL of endotoxin, yielding a level of sensitivity of 25 pg/mL when testing 5% HSA at a 1/10 dilution for the presence of endotoxin. Three batches of a therapeutic HSA that caused fever in humans were positive in the MM6/IL-6 method, whereas the same substances were negative in the BET and the RPT (**Table 9-8**). As in the BET, the samples required a 1/10 dilution to remove interfering substances. The assay had sensitivity equal to that of the BET (25 pg/mL) and 40-fold greater than the RPT (1000 pg/mL). The authors suggest that the MM6/IL-6 method represents an important alternative to the existing pyrogen tests and may be a more appropriate end-product test for the detection of pyrogens in parenteral products, such as HSA, that cannot be detected in the BET.

Table 9-8 Results of Pyrogen Testing of Batches of Therapeutic HSA Using the MM6/IL-6, BET, and RPT[1]

Batch of HSA	Endotoxin Quantitation by IL-6 Release (pg/mL)	Endotoxin Quantitation by the BET		RPT Result
		IU/mL	pg/mL[6]	
1[2]	97±2.3[3,4]	1.0-2.0	140-280	Pass
2[2]	30±2.8[4]	2.4-3.2	336-448	Pass
3[2]	31±2.3[4]	0.5-0.75	70-105	Pass
4	<25[5]	<0.24	<34	Pass
5	<25[5]	3.6-4.8[7]	504-762	Pass
6	<25[5]	<0.26	<36	Pass

Abbreviations: BET = Bacterial endotoxin test; HSA = Human serum albumin; IL-6 = Interleukin-6; IU = International units; MM6 = Mono Mac 6; RPT = Rabbit pyrogen test

[1]From Taktak et al. (1991)

[2]Batch of HSA used that caused fever in humans.

[3]Mean ± standard error of the mean

[4]Values are significantly different from subthreshold concentrations of endotoxin (<2.5 pg/mL; $p<0.001$).

[5]Values below the detection limit of the test system (25.0 endotoxin units/mL); preparations of HSA were tested at a dilution of 1/10, and 2.5 pg/mL endotoxin was the lowest concentration of endotoxin tested that evoked a significant release of IL-6.

[6]1.0 IU=0.14 ng for preparation used.

[7]False positive relative to RPT.

9.2 Conclusions from Scientific Literature Based on Independent Peer-Reviewed Reports and/or Reviews

An additional nine reports describing studies of cell-based *in vitro* pyrogen methods were obtained from the literature search described in **Section 1.5**. Although these reports did not include data on test substances that could be used in the performance analysis in **Section 6.0** and **Section 7.0**, they did evaluate the use of the *in vitro* pyrogen test methods for sensitivity to endotoxin (i.e., endotoxin detection limit), specificity of the response to endotoxin and/or non-endotoxin pyrogens (i.e., spectrum and relative potency of various pyrogens detected), and/or the impact of interfering substances. However, they did not compare results from the *in vitro* test methods to results from the RPT, BET, or human fever reaction. A summary of each study is presented below.

9.2.1 De Groote et al. (1992)

The authors measured the release of various cytokines (IL-1β, IL-6, TNF-α, IL-2, IFN- γ , and granulocyte-macrophage colony stimulating factor [GM-CSF]) in response to endotoxin or phytohemagglutinin (PHA) stimulation of WB and PBMC cultures. Endotoxin stimulated IL-1β, TNF-α, and IL-6 release, while PHA stimulated IL-2, IFN- γ , and GM-CSF release. There was a significant correlation between production of the three endotoxin-induced cytokines and the number of monocytes in the challenged culture, suggesting that monocytes are the major source of these cytokines: the other cytokines did not correlate with any of the cell types. The data also suggested that WB produced less variable levels of cytokines than PBMC on exposure to endotoxin. Consistent results were obtained with the WB test using more than 50 different blood donors. The authors suggest that WB is a more appropriate choice for studying cytokine production *in vitro* and its modulation by exogenous or endogenous factors, because natural cell-to-cell interactions are preserved, immune mediators are available, and cytokine levels obtained with PBMC were more variable.

9.2.2 Fennrich et al. (1999)

Fennrich and colleagues compared a commercially available human WB/IL-1β pyrogen assay (PyroCheck® from DPC Biermann, Bad Nauheim distributed by Millenia, U.K.) to the BET and RPT. There was a concentration-dependent IL-1β release in WB that was incubated with nitrocellulose filters containing live *E. coli* bacteria and *E. coli* killed by heat or by antibiotics. The authors also tested air conditioning filters from a veterinary sheep facility and identified filters to be contaminated with bacteria that were later confirmed by microbiological tests (the data and the identity of the organisms identified were not presented). The authors compared the PyroCheck®, BET, and the RPT (**Table 9-9**) and concluded that PyroCheck® is a simple, accurate test that detects a wider range of pyrogens than the BET.

Table 9-9 Comparison of the Application Spectra of the RPT, the BET, and the Human WB Assay (PyroCheck®)[1]

Test		Applications		
		PyroCheck®	BET	RPT
Pyrogens	Gram-negative	+	+	+
	Gram-positive	+	-	+
	Fungi	+	-	+
Product pyrogenicity	Biologicals	+	-	+
	Pharmaceuticals	+	+	+
	Medical devices	+	+[2]	-
	Air quality	+[2]	+[2]	-
	Blood products	+	-	-

Abbreviations: BET = Bacterial endotoxin test; RPT = Rabbit pyrogen test; WB = Whole blood
[1]From Fennrich et al. (1999)
[2]Based on preliminary data

9.2.3 Hansen and Christensen (1990)

This study compared the results from PBMC exposed to endotoxin or ultraviolet light-killed *S. aureus* as an index of pyrogenicity, and then compared these results to the BET and the RPT. The authors used human PBMC obtained from heparinized peripheral blood and measured IL-1-like material in culture supernatants by evaluating co-mitogenic activity on PHA-stimulated murine thymocytes (measured in units of IL-1β where 1 unit is defined as the concentration that gives 50% of the maximal incorporation of ^{3}H-thymidine in the thymocyte assay). The endpoint is referred to as an IL-1-like material because other cytokines such as IL-2, IL-6, and TNF-α may also stimulate the proliferative response of the thymocytes. When exposed to endotoxin, PBMC secreted cytokines in a concentration-dependent manner that provided a limit of detection of 200 pg/mL of endotoxin. In comparison, the BET can normally detect 10 to 100 pg/mL of endotoxin, while the RPT can detect 500 pg/mL. Therefore, the PBMC procedure had a level of detection of endotoxin 2.5-fold lower than that of the RPT and 2-fold higher than the BET. The PBMCs also responded with greater sensitivity to the Gram-positive pyrogen *S. aureus* (10^5 cells/mL), which was not detected in the BET (10^9 cells/mL). Based on these results, the authors proposed that the PBMC test be used as an alternative *in vitro* test to the BET and RPT.

9.2.4 Hartung and Wendel (1996)

The authors stimulated human WB with various inflammatory agents to release endogenous cytokines (i.e., IL-1β, TNF-α) and inflammatory mediators (i.e., prostaglandin E_2) as an *in vitro* method for the detection of pyrogenic materials. Cytokines were released in a concentration-dependent manner following exposure to endotoxin or LTA. Heat-killed Gram-positive bacteria (*S. aureus*) or components of these organisms (i.e., muropeptides, LTA, enterotoxins, streptolysin O) and plant mitogens such as phorbol myristate acetate and PHA also produced a cytokine response. Higher concentrations (three orders of magnitude) of the Gram-positive pyrogens were needed to elicit a response as compared to Gram-negative pyrogenic material.

Studies to determine the variability among the responses of different donor WB samples were also performed. Only two of the 18 donor samples released IL-1β in response to 1 pg/mL of endotoxin, but all responded to 10 pg/mL endotoxin. The release of IL-1β from the WB samples of 45 individual donors exposed to 100 ng/mL of endotoxin was also consistent. Based on these results, the authors suggested using the WB/IL-1β test method as an *in vitro* alternative to the RPT.

9.2.5 Moesby et al. (1999)

Moesby and colleagues compared pyrogen testing using MM6 cells, isolated PBMC, and the BET. LPS and ultraviolet (UV) light-killed Gram-negative *Staphylococcus typhimurium* or Gram-positive *S. aureus* produced concentration-dependent increases in IL-6 production in MM6 or PBMC cultures. PBMC, but not MM6 cells, were able to differentiate UV-irradiated yeast (*C. albicans*) and mold (*Aspergillus niger*) pyrogens, as evidenced by statistically significant increases in IL-6 production. The BET can detect Gram-negative endotoxin, but not Gram-positive endotoxin or LTA (the pyrogenic component of Gram-positive bacteria),

and it may weakly detect yeast or viral pyrogens that the MM6 assay could not detect. Therefore, the authors suggest that pyrogen testing using MM6 cells would be a useful supplement to the BET for the detection of both Gram-negative and Gram-positive bacteria.

9.2.6 Nakagawa et al. (2002)

Nakagawa and colleagues describe an *in vitro* pyrogen test system based on proinflammatory cytokine release from a sub-clone of MM6 cells (i.e., MM6-CA8) and compare this response to a human WB culture system and the RPT. Similar to MM6 cells, MM6-CA8 were developed for superior reactivity to both endotoxin and PG. The MM6-CA8 cells release IL-6, TNF-α, and IL-1, but in greater quantities than MM6 cells in the range of 1 to 1000 pg/mL of endotoxin (up to 4-fold greater) or to 1 to 1000 ng/mL PG (up to 10-fold greater) compared to MM6 cells. The range of responses of human WB to the various pyrogens was similar to that of the MM6-CA8 cells. The relative potencies of the various pyrogens in the RPT were similar to those of the cytokine-induction potencies in the WB and MM6-CA8 methods, except for polyinosinic:polycytidylic acid, which was reported to be 10,000-fold more potent as a pyrogen injected in rabbits when compared to humans. The authors conclude that these results suggest MM6-CA8 cells can detect a variety of pyrogens using IL-6 as the marker, and that these responses are highly relevant to the prediction of human fever reactions.

9.2.7 Pool et al. (1999)

This article describes a method to differentiate between endotoxin and non-endotoxin pyrogens when testing HSA solutions in a WB culture assay. Detection limits for four Gram-positive (*Bacillus stearothermophilus*, *B. subtilis*, *Micrococcus luteus*, and *S. aureus*) and four Gram-negative bacteria (*E. coli*, *Kleibsiella pneumoniae*, *Proteus vulgaris*, and *Pseudomonas aeruginosa*) were expressed as the number of whole bacteria required to produce a pyrogenic response equal to that of 1.25 EU/mL endotoxin. *B. stearothermophilus* and *E. coli* produced concentration-dependent increases in IL-6 production. The cationic antibiotic Polymyxin B, which inhibits the binding of endotoxin to the CD14 receptor, produced concentration-dependent inhibition of IL-6 release following exposure to 10 EU/mL endotoxin in the WB assay at concentrations up to 1 EU/mL and completely inhibited IL-6 release at concentrations above 2 EU/mL. In contrast, Polymyxin B had no effect on IL-6 release following exposure to *B. subtilis*. These data suggest that Polymyxin B may be useful for differentiating endotoxin and non-endotoxin pyrogenic contaminants. The data also suggests that binding of endotoxin to Polymyxin B (e.g., by linkage to an affinity column) may be used in the depyrogenation process.

9.2.8 Poole et al. (2003)

This paper describes a rapid single-plate *in vitro* test for the presence of pyrogenic substances based on monocyte activation. The assay uses polyclonal antibodies to IL-6 or TNF-α cytokines, coated and stabilized onto 96-well plates. Monocytoid cells (e.g., PBMC, MM6 or THP-1 cells), endotoxin standard (LPS), test sample, and a second biotinylated antibody specific for the cytokine (e.g., either IL-6 or TNF-α) are incubated for 2 to 4 hr in the antibody-coated wells. An ELISA for one of the cytokines is then performed on the washed plate. IL-6 is preferred and provides a limit of detection of 0.015 EU/mL with PBMC, 0.05

EU/mL in MM6 cells, and 0.03 EU/mL with diluted WB. The amount of TNF-α released in WB in response to endotoxin was approximately 50 to 70% lower than IL-6, but was released earlier (i.e., 2 vs. 4 hr). The amount of IL-6 released on exposure to endotoxin tended to be greater in this single plate test when compared to the traditional two-plate test (i.e., in which the supernatant from one plate is transferred to a second plate for the ELISA) using PBMCs, MM6 cells, THP-1 cells, or WB. The authors report that this single plate assay using IL-6 release as the endpoint can be completed in 5 hr, and that this time could be reduced to 3 hr using TNF-α as the endpoint (because it is released earlier from the cells). The authors also suggest that this single plate test method is readily adaptable to high-throughput assays.

9.2.9 Schindler et al. (2004)

The authors optimized conditions for use of cryopreserved human WB in pyrogen testing to obviate the need for fresh WB. The release of IL-1β from fresh and Cryo WB collected from five donors was used as the measure of endotoxin presence. Challenge with 0.5 or 1.0 EU/mL endotoxin resulted in IL-1β release in bloods from all donors, although kinetic studies suggested that IL-1β release was delayed one hr in the cryopreserved samples. Cryopreservation did not appear to alter the spectrum of detectable pyrogens or immune stimuli when results were compared to that of fresh WB, and no cytokine release was measured in materials that fresh WB did not respond to. Seven clinical-grade (i.e., endotoxin-free) parenteral products spiked with 0.5 EU/mL of endotoxin revealed that there was less interference in Cryo WB than in fresh WB based on lower minimal interference dilutions that were always at or below the MVD for each product. The data showed that a broad variety of drugs could be tested for pyrogenic contaminants using Cryo WB while maintaining the ELC established in the various Pharmacopeias.

[This Page Intentionally Left Blank]

10.0 Animal Welfare Considerations (Refinement, Reduction, and Replacement)

10.1 How the Five *In Vitro* Test Methods Will Refine, Reduce, or Replace Animal Use

ICCVAM promotes the scientific validation and regulatory acceptance of new methods that refine, reduce, or replace animal use where scientifically feasible. Refinement, Reduction, and Replacement are known as the three "Rs" of animal alternatives. These principles of humane treatment of laboratory animals are described as:

- Refining experimental procedures such that animal suffering is minimized
- Reducing animal use through improved science and experimental design
- Replacing animal models with non-animal procedures (e.g., *in vitro* technologies), where possible (Russell and Burch 1959)

In 2002, a total of 243,838 rabbits were used in the U.S. for all research and testing purposes, of which 6,324 rabbits were reported as experiencing more than slight or momentary pain and/or distress where anesthetics, analgesics, or tranquilizers could not be administered for scientific reasons (USDA 2002). Eight of these cases were specifically attributed to pyrogenicity testing, presumably based on induction of a fever response (USDA 2002). Thus, although the potential for more than slight or momentary pain and/or distress exists for pyrogenicity testing when a fever response is induced, it does not appear that a fever response is common. In 2006, a total of 239,720 rabbits were used in the U.S. for all research and testing purposes (USDA 2006). No data related to pyrogenicity testing were reported.

In Canada, a total of 18,152 rabbits were used for all scientific purposes in 2006, 3,485 of which were used for regulatory studies and the development of products (Canadian Council on Animal Care [CCAC] 2007). Although no specific data for the number of animals used for pyrogenicity testing were reported, it is likely that the number of rabbits used for this purpose is less than the total of 3,485 used for both regulatory studies and product development.

In the EU, approximately 313,000 total rabbits were used for all scientific purposes in 2005 (Commission of the European Communites [CEC] 2007). Of these, approximately 276,000 rabbits were used for pharmaceutical products and medical device testing (i.e., either research and development, production and quality control, or toxicological and other safety evaluations). Although the number of rabbits specifically used for pyrogenicity testing was not reported, it is likely that this number is significantly less than the total of 276,000.

In the U.K., a total of 21,736 procedures (which used 14,712 total rabbits due to reuse of some test animals) were performed using rabbits for all scientific purposes in 2004 (Home Office 2005). Of these procedures, 8,488 were specifically attributed to pyrogenicity testing in rabbits. Although the total number of rabbits used for these procedures were not provided, it is likely less than 8,488 rabbits based on the assumption that some animals were reused. In 2006, a total of 20,378 procedures (which included 13,397 total rabbits) were performed in the U.K. for all scientific purposes (Home Office 2007). No specific data for pyrogenicity testing were reported in 2006.

The currently accepted pyrogen test methods require the use of rabbits or horseshoe crab hemolymph. The proposed *in vitro* pyrogen test methods address each aspect of animal welfare outlined above. These assays use monocytoid cells of human origin, obtained either from WB donations or from an immortalized cell line. The capability of these five *in vitro* assays to detect Gram-negative endotoxin suggests that they may reduce or eventually replace the use of rabbits and/or horseshoe crab hemolymph for pyrogen testing. However, at the present time, the RPT detects classes of pyrogens that have neither been examined nor validated with the *in vitro* pyrogen test methods and thus, the RPT will still be required for most test substances.

10.2 Requirement for the Use of Animals

10.2.1 Rationale for the Use of Animals

Human blood donations are required for four of the five *in vitro* test methods (WB/IL-1β, WB/IL-6, Cryo WB/IL-1β, and PBMC/IL-6) proposed as replacements for the RPT, and as such, humans are the animals used for these assays. While the collection of human blood is a common medical procedure, the many aspects of human blood collection must be considered to ensure that human donors are appropriately treated.

11.0 Practical Considerations

Several issues are taken into account when assessing the practicality of using an *in vitro* test method in place of an *in vivo* test method. In addition to reliability and accuracy evaluations, assessments of the laboratory equipment and supplies needed to conduct the *in vitro* test method, level of personnel training, labor costs, and the time required to complete the test method relative to the *in vivo* test method are necessary. The time, personnel cost, and effort required to conduct the proposed test method(s) must be considered to be reasonable when compared to the *in vivo* test method it is intended to replace.

11.1 Transferability of the *In Vitro* Pyrogen Test Methods

Test method transferability addresses the ability of a method to be accurately and reliably performed by multiple laboratories (ICCVAM 2003), including those experienced in the particular type of procedure as well as laboratories with less or no experience in the particular procedure. The degree of transferability of a test method can be evaluated by its interlaboratory reproducibility. ECVAM measured the transferability (i.e., interlaboratory reproducibility) of each assay among experienced laboratories. The results presented in **Tables 7-3** and **7-4** provide an estimate of the minimum variability to be expected. Interlaboratory variability is anticipated to be greater (i.e., lower transferability) among laboratories that have less experience with the assays.

11.1.1 Facilities and Major Fixed Equipment

A standard laboratory facility for sterile tissue culture is necessary for performing the *in vitro* pyrogen test methods. The major equipment necessary to conduct the tests are readily available and include a laminar flow hood, tissue culture incubator, water bath, and spectrophotometric microplate reader.

In contrast, the RPT requires a facility that meets applicable State and Federal regulations for the care and housing of laboratory animals. The primary expense for equipping a facility to conduct the RPT would be the acquisition of an adequate animal room and associated housing (e.g., cages, bedding, food, water, etc.) for boarding animals during the study.

11.1.2 General Availability of Other Necessary Equipment and Supplies

The equipment and supplies necessary to conduct the *in vitro* pyrogen test methods (e.g., micropipetters, sterile tissue culture vessels, disposable plastic ware, assay reagents) are readily available in most scientific laboratories, or can be obtained from any of several scientific laboratory equipment vendors.

The RPT requires fewer general laboratory supplies. Those that are needed are readily available in most laboratories, or could be readily obtained from any of a number of scientific laboratory equipment vendors.

11.2 Personnel Training Considerations

Training considerations are defined as the level of instruction needed for personnel to conduct the test method accurately and reliably (ICCVAM 2003). Evaluation of the levels of

training and expertise needed to conduct the test method, as well as the training requirements needed to insure that personnel are competent in the test procedures, are discussed below.

11.2.1 Required Training and Expertise Needed to Conduct the In Vitro Pyrogen Test Methods

Laboratory personnel require training with the relevant ELISA procedures and the aseptic techniques associated with mammalian tissue culture. The quality criteria associated with each *in vitro* test method may be used to ensure that personnel are competent in the performance of the various procedures. When a technician has mastered all aspects of the protocol, and can independently conduct the assay such that the quality criteria have been met, the individual is considered to have demonstrated proficiency in the assay.

The RPT requires training in the care and handling of laboratory animals, and the collection of accurate rectal temperature measurements at the appropriate time intervals from each rabbit. The laboratory personnel must be adequately trained to maintain the animals, and to accurately and consistently record the proper body temperature. It is not known what, if any, proficiency requirements are in place for the RPT.

11.3 Cost Considerations

In addition to the major fixed equipment and overhead requirements, three additional factors contribute to the overall cost of the proposed *in vitro* test methods: 1) cost and licensing fees associated with the MM6 monocytoid cell line, 2) cost of the reagents for the ELISA procedure, and 3) personnel costs associated with obtaining human blood and performing the test methods. With respect to the RPT, the direct and indirect costs of operating an animal facility must be considered. The most notable expenses will likely include personnel to care for the maintenance of the rabbits, staff to perform the RPT, and veterinarians to monitor the health of the rabbits. As summarized in **Table 11-1**, cost estimates from various contract laboratories that perform the RPT or from one contract laboratory that performs an ELISA-based *in vitro* pyrogen test using human WB indicate that the *in vitro* test methods are considerably more cost effective (i.e., by about a factor of ten) than the RPT. Furthermore, the use of high throughput procedures to analyze the *in vitro* pyrogen tests may provide further reduced costs per test substance.

11.4 Time Considerations

The *in vitro* pyrogen methods require two half-days (i.e., one before and one after the overnight incubation) to complete if cryopreserved blood or MM6 cells are available. If fresh WB is used or if interference testing is needed, additional time will be required. On the first day, the test materials are prepared and incubated with the monocytoid cells. On the second day, cytokine release from the cells is determined by an ELISA procedure. The BET and RPT can both be completed within one working day. However, according to the USP30 NF25<151> (USP 2007b) procedure for the RPT, each rabbit must be conditioned prior to its first use by a sham test that includes all steps of pyrogenicity testing except for injection.

Table 11-1 Cost Estimates for the RPT and *In Vitro* Pyrogen Tests

Contract Laboratory	Test or Cell Line	GLP Compliant	Cost Estimate per Test	Additional Information
A	RPT	Yes	$2100[1]	-
B	RPT	Yes	$4050[1]	-
C	RPT	Yes	$3600[1]	-
D	IPT/HumanWB	ND	$315[2]	Cost decreases with number of test substances; $315 per 1 test substance; $210 per 2 to 10 test substances; $105 per 11 or more test substances. Note: IPT is not a licensed product and should not be used for the release of drugs.
E	MM6	NA	Negotiable	Use of MM6 cells for product testing require negotiation of a fee for provision and a royalty payment per batch of product tested with Dr. HWL Ziegler-Heitbrock at the University of Leicester, Dept of Microbiology, Leicester, U.K.

Abbreviations: GLP = Good laboratory practice; IPT = *In vitro* pyrogen test; MM6 = Mono Mac 6; NA = Not applicable; ND = Not determined; RPT = Rabbit pyrogen test; WB = Whole blood

[1]Each RPT includes one test substance, one positive, and one negative control performed in triplicate. Thus, a minimum of 9 rabbits is needed per test.

[2]Each IPT includes one test substance, one positive, and one negative control performed in triplicate.

[This Page Intentionally Left Blank]

12.0 References

Andrade SS, Silveira RL, Schmidt CA, Junior LB, Dalmora SL. 2003. Comparative evaluation of the human whole blood and human peripheral blood monocyte tests for pyrogens. Int J Pharm 265:115-124.

Barnett V, Lewis T. 1984. Outliers in statistical data. In: Wiley Series in Probability and Mathematical Statistics. (Barnett V, Lewis T, eds). 3rd ed. New York:John Wiley and Sons.

Beutler B, Rietschel ET. 2003. Innate immune sensing and its roots: the story of endotoxin. Nat Rev Immunol 3:169-176.

Bleeker WK, de Groot EM, den Boer PJ, Biessels PT, Asrden LA, Bakker JC. 1994. Measurement of interleukin-6 production by monocytes for *in vitro* safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22:835-840.

Brunson KW, Watson DW. 1974. Pyrogenic specificity of streptococcal exotoxins, staphylococcal enterotoxin, and gram-negative endotoxin. Infect Immun 10:347-351.

Burrell R. 1994. Human responses to bacterial endotoxin. Circ Shock 43:137-153.

Carlin G, Viitanen E. 2003. *In vitro* pyrogenicity of a multivalent vaccine; infanrix. Pharmeuropa 15:418-423.

Carlin G, Viitanen E. 2005. *In vitro* pyrogenicity of the diphtheria, tetanus and acellular pertussis components of a trivalent vaccine. Vaccine 23:3709-3715.

CCAC. 2007. CCAC Survey of Animal Use 2006. Ottawa, Ontario:Canadian Council on Animal Care (CCAC). Available: http://www.ccac.ca.

CEC. 2007. Report from the Commission to the Council and the European Parliament. Fifth Report on the Statistics on the Number of Animals Used for Experimental and Other Purposes in the Member States of the European Union (Annex). Brussels:Commission of the European Communities (CEC). Available: http://ec.europa.eu/environment/chemicals/lab_animals/pdf/5th_stat_rep_lab_animals_en.pdf

Conti B, Tabarean I, Andrei C, Bartfai T. 2004. Cytokines and fever. Front Biosci 9:1433-1449.

Cooper JF, Levin J, Wagner HN. 1971. Quantitative comparison of *in vitro* and *in vivo* methods for the detection of endotoxin. J Lab Clin Med 78:138-148.

Co Tui, Shrift, MH. 1942. Alternative test for pyrogens in infusion fluids. Proc Soc Exp Biol Med 49:320-330.

Daneshian M, Guenther A, Wendel A, Hartung T, von Aulock S. 2006. *In vitro* pyrogen test for toxin or immunomodulatory drugs. J Immunol Methods 313:169-175.

De Groote D, Zangerle PF, Gevaert Y, Fassotte MF, Beguin Y, Noizat-Pirenne F, et al. 1992. Direct stimulation of cytokines (IL-1β, TNF-α, IL-6, IL-2, IFN- γ and GM-CSF) in whole blood. I. Comparison with isolated PBMC stimulation. Cytokine 4:239-248.

Dinarello CA, O'Conner JV, LoPreste G, Swift RL. 1984. Human leukocytic pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia Coli. J Clin Microbiol 20:323-329.

Dinarello CA. 1999. Cytokines as endogenous pyrogens. J Infect Dis 179:S294-304.

Dinarello CA. 2004. Infection, fever, and exogenous and endogenous pyrogens: some concepts have changed. J Endotoxin Res 10:201-222.

Ding JL, Chai C, Pui AWM, Ho B. 1997. Expression and full length and deletion homologues of *Carcinoscorpus rotundicauda* Factor C in *Saccharomyces cerevisciae*: immunoreactivity and endotoxin binding. J Endotoxin Res 4:33-43.

Ding JL, Ho B. 1998. Cloned Factor C cDNA of the Singapore horseshoe crab *Carsinoscorpius rotundicauda* and purification of Factor C proenzyme. U.S. Patent No. 5 716 834.

Ding JL, Ho B. 2001. A new era in pyrogen testing. Trends Biotechnol 19:277-281.

Dixon WJ. 1950. Analysis of extreme values. Ann Math Stat 21:488-506.

Duff GW, Atkins E. 1982. The detection of endotoxin by *in vitro* production of endogenous pyrogen: comparison with *Limulus* amebocyte lysate gelation. J Immunol Methods 52:323-331.

EP. 2005a. Biological Tests. Supplement 5.2. 2.6.8 Pyrogens. In: European Pharmacopoeia (EP). 5th Ed. Strasbourg, France:Council of Europe - European Directorate for the Quality of Medicines.

EP. 2005b. Biological Tests. Supplement 5.2. 2.6.14 Bacterial Endotoxins. In: European Phamacopoeia (EP). 5th Ed. Strasbourg, France:Council of Europe - European Directorate for the Quality of Medicines.

EPA. 2003a. Good Laboratory Practice Standards. Toxic Substances Control Act. 40 CFR 792.

EPA. 2003b. Good Laboratory Practice Standards. Federal Insecticide, Fungicide, and Rodenticide Act. 40 CFR 160.

Eperon S, Jungi TW. 1996. The use of human monocytoid lines as indicators of endotoxin. J Immunol Methods 194(2):121-129.

Eperon S, de Groote DD, Werner-Felmayer G, Jungi TW. 1997. Human monocytoid cell lines as indicators of endotoxin: comparison with rabbit pyrogen and *Limulus* amoebocyte lysate assay. J Immunol Methods 207:135-145.

FDA. 2003. Good Laboratory Practices for Nonclinical Laboratory Studies. 21 CFR 58.

FDA. 2005. General Biological Products Standards: General Provisions: Purity. 21 CFR 610.13.

Fennrich S, Wendel A, Hartung T. 1999. New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 16:146-149.

Francois C, Neveu J, Sauvaire D, Bonnet PA, Tissier MH. 2006. *In vitro* pyrogenicity of Gram-positive bacteria - validation of the kit using fresh human whole blood. Pharmeur Sci Notes 2006:17-21.

Gaines Das RE, Brügger P, Patel M, Mistry Y, Poole S. 2004. Monocyte activation test for pro-inflammatory and pyrogenic contaminants of parenteral drugs: test design and data analysis. J. Immunol Methods 288:165-177.

Greisman SE, Hornick RB. 1969. Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131:1154-1158.

Grubbs FE. 1969. Procedures for detecting outlying observations in samples. Technometrics 11:1-21.

Hansen EW, Christensen JD. 1990. Comparison of cultured human mononuclear cells, *Limulus* amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther 15:425-433.

Hartung T, Wendel A. 1996. Detection of pyrogens using human whole blood. In Vitro Toxicol 9:353-359.

Hartung T, Aaberge I, Berthold S, Carlin G, Charton E, Coecke S, et al. 2001. Novel pyrogen tests based on the human fever reaction. The report and recommendations of ECVAM Workshop 43. Altern Lab Anim 29:99-123.

Hochstein HD, Fitzgerald EA, McMahon FG, Vargas R. 1994. Properties of US standard endotoxin (EC-5) in human male volunteers. J Endotoxin Res 1:52-56.

Hoffmann S, Peterbauer A, Schindler S, Fennrich S, Poole S, Mistry Y, et al. 2005a. International validation of novel pyrogen tests based on human monocytoid cells. J Immunol Methods 298:161-173.

Hoffmann S, Luderitz-Puchel U, Montag T, Hartung T. 2005b. Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modelling. J Endo Res 11:25-31.

Holzhütter HG, Archer G, Dami N, Lovell DP, Saltelli A, Sjostrom M. 1996. Recommendations of the application of biostatistical methods during the development and validation of alternative toxicological methods. ECVAM Biostatistics Task Force Report 1. Altern Lab Anim 24:511-530.

Home Office. 2005. Statistics of Scientific Procedures on Living Animals 2004. London:The Stationery Office. Available: http://www.official-documents.gov.uk/document/cm67/6713/6713.pdf

Home Office. 2007. Statistics of Scientific Procedures on Living Animals 2006. London:The Stationery Office. Available: http://www.homeoffice.gov.uk/rds/pdfs07/spanimals06.pdf

ICCVAM. 1997. Validation and Regulatory Acceptance of Toxicological Test Methods: A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods. NIH Publication No: 97-3981. Research Triangle Park, NC:National Toxicology Program.

ICCVAM. 2000. ICCVAM Authorization Act of 2000. Public Law 106-545 (42 U.S.C. 285).

ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No: 03-4508. Research Triangle Park, NC:National Toxicology Program.

Ishii KJ, Coban C, Akira S. 2005. Manifold mechanisms of toll-like receptor-ligand recognition. 2005. J Clin Immunol 25:511-521.

Ishii KJ, Akira S. Innate immune recognition of, and regulation by, DNA. 2006. Trends Immunol. 27:525-532.

ISO. 2005. General Requirements for the Competence of Testing and Calibration Laboratories (Second edition). ISO/IEC 17025. Geneva:International Organization for Standardization (ISO).

ISO. 2006. Biological Evaluation of Medical Devices - Part 11: Tests for Systemic Toxicity (First edition 1993-12-15). ISO 10993-11. Geneva:International Organization for Standardization (ISO).

JP. 2001. Bacterial Endotoxins Test. In: The Japanese Pharmacopoeia 14th ed. (JP XIV) Part I. Shibuya, Shiubya-ku, Tokyo, Japan:The Society of Japanese Pharmacopoeia.

Keene WR, Silberman HR, Landy M. 1961. Observations on the pyrogenic response to and its application to the bioassay of endotoxin. J Clin Invest 40:295-301.

Marth E. and Kleinhappl B. 2002. Albumin is a necessary stabilizer of TBE-vaccine to avoid fever in children after vaccination. Vaccine 20:532-537.

Martich GD, Boujoukos AJ, Suffredini AF. 1993. Response of man to endotoxin. Immunobiology 187:403-416.

Martin MA, Roberts S. 2006. An evaluation of bootstrap methods for outlier detection in least squares regression. J Appl Stat 33:703-720.

Martis L, Patel M, Giertych J, Mongoven J, Tammine M, Perrier MA, et al. 2005. Aseptic perontonitis due to peptidoglycan contamination of pharmacopoeia standard dialysis solution. Lancet 365:588-594.

McKinney BA, Reif DM, Rock MT, Edwards KM, Kingsmore SF, Moore JH, Crowe Jr JE. 2006. Cytokine expression patterns associated with systemic adverse events following smallpox immunization. J Infect Dis 194:444-453.

Moesby L, Jensen S, Hansen EW, Christensen JD. 1999. A comparative study of Mono Mac 6 cells, isolated mononuclear cells and *Limulus* amoebocyte lysate assay in pyrogen testing. Int J Pharm 191:141-149.

Moesby L, Hansen EW, Christensen JD. 2000. Endotoxin testing of proteins for parenteral administration using the Mono Mac 6 assay. J Clin Pharm Ther 25:283.

Moesby L, Hansen EW, Christensen JD, Høyer CH, Juhl GL, Olsen HB. 2005. Dry and moist heat sterilization cannot inactivate pyrogenicity of Gram positive microorganisms. Eur J Pharm Sci 26:318-323.

Nakagawa Y, Maeda H, Murai T. 2002. Evaluation of the *in vitro* pyrogen test system based on proinflammatory cytokine release from human monocytes: comparison with a human

whole blood culture test system and with the rabbit pyrogen test. Clin Diagn Lab Immunol 9:588-597.

Netea MG, Kallberg BJ, Van der Meer JWM. 2000. Circulating cytokines as mediators of fever. Clin Inf Dis 31:S178-S184.

Norata GD, Garlaschelli K, Ongari M, Raselli S, Gigore L, Benenuto F, Maggi FM, Catapano AL. 2005. Effect of the Toll-like receptor 4 (TLR-4) variants on intima-media thickness and monocyte-derived macrophage response to LPS. J Intern Med 258:21-27.

OECD. 1998. OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring Number 1: OECD Principles on Good Laboratory Practice (as revised in 1997). ENV/MC/CHEM(98)17. Paris:Organisation for Economic Co-operation and Development (OECD).

Park JM, Ng VH, Maeda S, Rest RF, Karin M. 2004. Anthrolysin O and other Gram-positive cytolysins are Toll-like receptor 4 agonists. J Exp Med 200:1647-1655.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. 1998. The detection of pyrogens in blood products using an ex vivo whole blood culture assay. J Immunoassay 19:95-111.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. 1999. Differentiation between endotoxin and non-endotoxin pyrogens in human albumin solutions using an *ex vivo* whole blood culture assay. J Immunoassay 20:79-89.

Poole S, Mistry Y, Ball C, Gaines Das RE, Opie LP, Tucker G, et al. 2003. A rapid 'one-plate' *in vitro* test for pyrogens. J Immunol Methods 274:209-220.

Poole S, Thorpe R, Meager A, Gearing AJ. 1988. Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69:121-123.

Russell WMS, Burch RL. 1959. The Principles of Humane Experimental Technique. Potters Bar, Herts, England:Universities Federation for Animal Welfare.

Rylander R. 2002. Endotoxin in the environment—exposure and effects. J Endotoxin Res 8:241-252.

Saper CB, Breder CD. 1994. The neurological basis of fever. New England J Med 330:1880-1886.

Schindler S, Asmus S, von Aulock S, Wendel A, Hartung T, Fennrich S. 2004. Cyropreservation of human whole blood for pyrogencity testing. J Immunol Methods 294:89-100.

Schindler S, Spreitzer I, Löschner B, Hoffmann S, Hennes K, Halder M, et al. 2006. International validation of pyrogen tests based on cryopreserved human primary blood cells. J Immunol Methods 316:42-51.

Stoll LL, Denning GM, Weintraub NL. 2006. Endotoxin, TLR4 signalling and vascular inflammation: potential therapeutic targets in cardiovascular disease. 12:4229-4245.

Taktak YS, Selkirk S, Bristow AF, Carpenter A, Ball C, Rafferty B, et al. 1991. Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43:578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y, Kobayashi Y, Konno T, Tada K. 1980. Establishment and characterization of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26:171-176.

USDA. 2002. Animals Used in Research 2002. Available: http://www.aphis.usda.gov/animal_welfare/downloads/awreports/awreport2002.pdf

USDA. 2006. Animals Used in Research 2006. Available: http://www.aphis.usda.gov/animal_welfare/downloads/awreports/awreport2006.pdf

USP. 2007a. Bacterial Endotoxins Test. USP30 NF25<85>. Rockville, MD:The U.S. Pharmacopeial (USP) Convention.

USP. 2007b. Pyrogen Test. USP30 NF25<151>. Rockville, MD:The U.S. Pharmacopeial (USP) Convention.

van Deventer SJH. 2000. Cytokine and cytokine receptor polymorphisms in infectious disease. Intensive Care Med 26:S98-S102.

von Aulock S, Schroder NWJ, Gueninzius K, Traub S, Hofffmann S, Graf K, Dimmeler S, Hartung T, Schumann RR, Hermann C. 2003. Heterozygous Toll-like receptor 4 polymorphism does not influence lipopolysaccharide-induced cytokine release in human whole blood. J Infect Dis 188:938-943.

Walls EA, Berkson J, Smith SA. 2002. The horseshoe crab, *Limulus polyphemus*: 200 million years of existence, 100 years of study. Rev Fish Sci 10:39-73.

Westphal O. 1956. In: Polysaccharides in Biology (GF Springer, ed.). Trans. 2nd Conf. New York:Josiah Macy Jr. Foundation. 115.

Ziegler-Heitbrock HW, Thiel E, Futterer A, Herzog V, Wirtz A, Riethmuller G. 1988. Establishment of a human cell line (Mono Mac 6) with characteristics of mature monocytes. Int J Cancer 41:456-461.

13.0 Glossary[1]

Accuracy[2]: (a) The closeness of agreement between a test method result and an accepted reference value. (b) The proportion of correct outcomes of a test method. Accuracy is a meaure of test method performance and one aspect of relevance. The term is often used interchangeably with concordance (see two-by-two table). Accuracy is highly dependent on the prevalence of positives in the poulation being examined.

Amebocytes: The blood cells of the horseshoe crab (*Limulus polyphemus* or *Tachypleus tridentatus*) that contain the active components of the reagent used in the BET.

Assay[2]: The experimental system used. Often used interchangeably with "test" and "test method."

Bacterial endotoxin test (BET)[3]: A test used to quantify endotoxins of Gram-negative bacterial origin using amebocyte lysate from the horseshoe crab. Two types of techniques exist: the gel-clot techniques, which are based on gel formation and the photometric techniques. The photometric techniques include the turbidimetric technique, which is based on the development of turbidity after cleavage of an endogenous substrate and a chromogenic method, which is based on the development of color after cleavage of a synthetic peptide-chromogen complex.

Coded substances: Substances labeled by code rather than name so that they can be tested and evaluated without knowledge of their identity or anticipation of test results. Coded substances are used to avoid intentional or unintentional bias when evaluating laboratory or test method performance.

Coefficient of variation (CV): A statistical representation of the precision of a test. It is expressed as a percentage and is calculated as follows:

$$\left(\frac{standard\ deviation}{mean}\right) \times 100\%$$

Concordance[2]: The proportion of all substances tested that are correctly classified as positive or negative. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with accuracy (see two-by-two table). Concordance is highly dependent on the prevalence of positives in the population being examined.

Endogenous pyrogens: Various cytokines including interleukins (e.g., IL-1α, IL-1β), tumor necrosis factor (i.e., TNF-α, TNF-β), and interferon (IFN-γ) released from leukocytes in response to external stimuli (e.g., endotoxin) capable of causing an increase in body temperature above the normal level.

---

[1]The definitions in this Glossary are restricted to the RPT, the *in vitro* pyrogen test methods included in this BRD, and the BET.
[2]From ICCVAM (2003)
[3]From USP (2005)

Endotoxin limit concentration (ELC): The concentration at which endotoxin is considered to be pyrogenic. It is expressed as the ratio of the threshold pyrogen dose (K) and the RPT dose or the maximum human dose administered on a weight (kg) basis in 1 hr (M) defined as K/M. The ELC varies based on M.

- The U.S. Food and Drug Administration (FDA) ELC for non-intrathecal medical devices is 0.5 EU/mL.
- The FDA ELC for intrathecal medical devices is 0.06 EU/mL.

Endpoint[2]: The biological or chemical process, response, or effect assessed by a test method.

False negative[2]: A substance incorrectly identified as negative by a test method.

False negative rate[2]: The proportion of all positive substances falsely identified by a test method as negative (see two-by-two table). It is one indicator of test method accuracy.

False positive[2]: A substance incorrectly identified as positive by a test method.

False positive rate[2]: The proportion of all negative substances that are falsely identified by a test method as positive (see two-by-two table). It is one indicator of test method accuracy.

Fever: Elevation of body temperature above the normal level.

Good laboratory practices (GLP)[2]: Regulations promulgated by the FDA and the U.S. Environmental Protection Agency, principles and procedures adopted by the Organization for Economic Cooperation and Development, and Japanese authorities that describe record keeping and QA procedures for laboratory records that will be the basis for data submissions to national regulatory agencies.

Hazard[2]: The potential for an adverse health or ecological effect. A hazard potential occurs only if an exposure occurs that leads to the possibility of an adverse effect being manifested.

Interlaboratory reproducibility[2]: A measure of whether different qualified laboratories using the same protocol and test substances can produce qualitatively and quantitatively similar results. Interlaboratory reproducibility is determined during the prevalidation and validation processes and indicates the extent to which a test method can be transferred successfully among laboratories.

Intralaboratory repeatability[2]: The closeness of agreement between test results obtained within a single laboratory when the procedure is performed on the same substance under identical conditions within a given time period.

Intralaboratory reproducibility[2]: The first stage of validation; a determination of whether qualified people within the same laboratory can successfully replicate results using a specific test protocol at different times.

In vitro: In glass. Refers to assays that are carried out in an artificial system (e.g., in a test tube or petri-dish) and typically use single-cell organisms, cultured cells, cell-free extracts, or purified cellular components.

In vivo: In the living organism. Refers to assays performed in multi-cellular organisms.

Lipopolysaccharide (LPS): A complex of lipid and carbohydrate (i.e., endotoxin) released from the cell walls of Gram-negative organisms that is pyrogenic and capable of producing septic shock.

Lipoteichoic acid: A polyol phosphate polymer bearing a strong negative charge that is covalently linked to the peptidoglycan in Gram-positive bacteria. It is strongly antigenic, but is generally absent in Gram-negative bacteria. Therefore, it is considered the primary pyrogenic component of Gram-positive bacteria.

Minimum valid concentration (MVC): The concentration of a product when it is diluted to the MVD expressed as λM/K, where:

- **λ** = The sensitivity of the *Limulus* Amebocyte Lysate (LAL) reagent used expressed as EU/mL. The value varies with the method employed. For the gel-clot method, it is the labeled LAL sensitivity (EU/mL). For the chromogenic, turbidometric, or kinetic-turbidometric methods, it is the lowest point used in the standard curve.
- **M** = The maximum human dose for pyrogenicity administered on a weight basis (kg) in 1 hr, or the RPT dose (whichever is larger). It is one of the variables used to define the ELC defined as the ratio of K/M, where K is the threshold pyrogen dose in rabbits or humans.
- **K** = See threshold pyrogen dose.

Maximum valid dilution (MVD): When a U.S. Pharmacopeia (USP) ELC is defined, the MVD is the ratio of the product of the ELC and the product potency to the LAL reagent sensitivity (λ) expressed as ([ELC x Product Potency]/λ). If there is no official USP ELC defined, then the MVD is the ratio of the Product Potency/MVC.

Monocytoid cells: Cells obtained from peripheral blood or grown in culture that phenotypically resemble monocytes or macrophages.

Negative control: An untreated sample containing all components of a test system, except the test substance solvent, which is replaced with a known non-reactive material, such as water. This sample is processed with test substance-treated samples and other control samples to determine whether the solvent interacts with the test system.

Negative predictivity[2]: The proportion of correct negative responses among substances testing negative by a test method (see two-by-two table). It is one indicator of test method accuracy. Negative predictivity is a function of the sensitivity of the test method and the prevalence of negatives among the substances tested.

Parenteral: Introduction into the body by some means other than through the gastrointestinal tract; referring particularly to intravenous (i.v.), intramuscular, subcutaneous, or intrathecal injection.

Performance[2]: The accuracy and reliability characteristics of a test method (see accuracy and reliability).

pH: A measure of the acidity or alkalinity of a solution. A pH of 7.0 is neutral; higher pHs are alkaline, lower pHs are acidic.

Positive control: A sample containing all components of a test system and treated with a substance known to induce a positive response, which is processed with the test substance-treated and other control samples to demonstrate the sensitivity of each experiment and to allow for an assessment of variability in the conduct of the assay over time.

Positive predictivity[2]**:** The proportion of correct positive responses among substances testing positive by a test method (see two-by-two table). It is one indicator of test method accuracy. Positive predictivity is a function of the sensitivity of the test method and the prevalence of positives among the substances tested.

Prevalence[2]**:** The proportion of positives in the population of substances tested (see two-by-two table).

Protocol[2]**:** The precise, step-by-step description of a test method, including a list of all necessary reagents and criteria and procedures for evaluation of the test data.

Pyrogen: A substance that causes a rise in body temperature above normal or that produces a fever. Gram-negative, Gram-positive, and acid-fast bacteria, molds, viruses, and yeast and some of their cellular constituents are pyrogenic.

Quality assurance (QA)[2]**:** A management process by which adherence to laboratory testing standards, requirements, and record keeping procedures is assessed independently by individuals other than those performing the testing.

Rabbit pyrogen test (RPT)[3]**:** A test designed to limit to an acceptable level the risks of febrile reaction in the patient to the administration, by injection, or the product concerned. The test involves measuring the rise in temperature of rabbits following the i.v. injection of a test solution.

Reduction alternative[2]**:** A new or modified test method that reduces the number of animals required.

Reference test method[2]**:** The accepted *in vivo* test method used for regulatory purposes to evaluate the potential of a test substance to be hazardous to the species of interest.

Refinement alternative[2]**:** A new or modified test method that refines procedures to lessen or eliminate pain or distress in animals or enhances animal well-being.

Relevance[2]**:** The extent to which a test method correctly predicts or measures the biological effect of interest in humans or another species of interest. Relevance incorporates consideration of the accuracy or concordance of a test method.

Reliability[2]**:** A measure of the degree to which a test method can be performed reproducibly within and among laboratories over time. It is assessed by calculating intra- and inter-laboratory reproducibility and intralaboratory repeatability.

Replacement alternative[2]**:** A new or modified test method that replaces animals with non-animal systems or one animal species with a phylogenetically lower one (e.g., a mammal with an invertebrate).

Reproducibility[2]: The consistency of individual test results obtained in a single laboratory (intralaboratory reproducibility) or in different laboratories (interlaboratory reproducibility) using the same protocol and test substances (see intra- and inter-laboratory reproducibility).

Sensitivity[2]: The proportion of all positive substances that are classified correctly as positive in a test method. It is a measure of test method accuracy (see two-by-two table).

Specificity[2]: The proportion of all negative substances that are classified correctly as negative in a test method. It is a measure of test method accuracy (see two-by-two table).

Test[2]: The experimental system used; often used interchangeably with "test method" and "assay."

Test method[2]: A process or procedure used to obtain information on the characteristics of a substance or agent. Toxicological test methods generate information regarding the ability of a substance or agent to produce a specified biological effect under specified conditions. Used interchangeably with "test" and "assay" (see validated test method and reference test).

Test method component: Structural, functional, and procedural elements of a test method that are used to develop the test method protocol. These components include unique characteristics of the test method, critical procedural details, and quality control measures.

Threshold pyrogen dose: The dose level at which a product is considered to be pyrogenic or non-pyrogenic. It is one of the variables (K) used to calculate the ELC defined as K/M, where M is the RPT dose or the maximum human dose administered in 1 hr (whichever is larger).

- The threshold pyrogen dose for non-intrathecal use in rabbits and humans is 5.0 EU/kg.
- The threshold pyrogen dose for intrathecal use in rabbits and humans is 0.2 EU/kg.

Tiered testing: A testing strategy where all existing information on a test substance is reviewed, in a specified order, prior to *in vivo* testing. If the irritancy potential of a test substance can be assigned, based on the existing information, no additional testing is required. If the irritancy potential of a test substance cannot be assigned, based on the existing information, a step-wise animal testing procedure is performed until an unequivocal classification can be made.

Transferability[2]: The ability of a test method or procedure to be accurately and reliably performed in different, competent laboratories.

Two-by-two table[2]: The two-by-two table can be used for calculating accuracy (concordance) ([a+d]/[a+b+c+d]), negative predictivity (d/[c+d]), positive predictivity (a/[a+b]), prevalence ([a+c]/[a+b+c+d]), sensitivity (a/[a+c]), specificity (d/[b+d]), false positive rate (b/[b+d]), and false negative rate (c/[a+c]).

		NEW TEST OUTCOME		
		Positive	**Negative**	**Total**
Reference Test Outcome	**Positive**	a	c	a + c
	Negative	b	d	b + d
	Total	a + b	c + d	a + b + c + d

Validated test method[2]: An accepted test method for which validation studies have been completed to determine the relevance and reliability of this method for a specific proposed use.

Validation[2]: The process by which the reliability and relevance of a procedure are established for a specific purpose.

Weight of evidence (process): The strengths and weaknesses of a collection of information are used as the basis for a conclusion that may not be evident from the individual data.

Appendix A

ECVAM BRDs and Standard Operating Procedures

A1 The Human Whole Blood (WB)/Interleukin (IL)-1β *In Vitro* Pyrogen Test..... A-3
A2 The Human WB/IL-1β *In Vitro* Pyrogen Test: Application of Cryopreserved (Cryo) Human WB A-133
A3 The Human WB/IL-6 *In Vitro* Pyrogen Test A-237
A4 The Human Peripheral Blood Mononuclear Cell (PBMC)/IL-6 *In Vitro* Pyrogen Test..... A-335
A5 The Monocytoid Cell Line Mono Mac 6 (MM6)/IL-6 *In Vitro* Pyrogen Test A-459

[This Page Intentionally Left Blank]

Appendix A1

The Human Whole Blood (WB)/Interleukin (IL)-1β *In Vitro* Pyrogen Test

ECVAM Background Review Document (March 2006) A-5
ECVAM Standard Operating Procedure (February 2002) A-85
ECVAM Standard Operating Procedure for the Validation Phase (September 2002) A-107

[This Page Intentionally Left Blank]

THE HUMAN WHOLE BLOOD/IL-1 IN VITRO PYROGEN TEST (WB/IL-1)

BRD: WB/IL-1 March, 2006

Contents

1 RATIONALE FOR THE PROPOSED TEST METHOD 4

1.1 INTRODUCTION 4
1.2 REGULATORY RATIONALE AND APPLICABILITY 5
1.3 SCIENTIFIC BASIS FOR THE PROPOSED TEST METHOD 6

2 TEST METHOD PROTOCOL COMPONENTS 8

2.1 OVERVIEW OF TEST METHOD 8
2.2 RATIONAL FOR SELECTED TEST COMPONENTS 9
2.3 BASIS FOR SELECTION OF THIS TEST METHOD 13
2.4 PROPRIETARY COMPONENTS 13
2.5 REPLICATES 14
2.6 MODIFICATIONS APPLIED AFTER VALIDATION 14
2.7 DIFFERENCES WITH SIMILAR TEST METHODS 14

3 SUBSTANCES USED FOR VALIDATION 15

3.1 SELECTION OF SUBSTANCES USED 15
3.2 NUMBER OF SUBSTANCES 15
3.3 DESCRIPTION OF SUBSTANCES USED 16
3.4 SAMPLE CODING PROCEDURE 17
3.5 RECOMMENDED REFERENCE CHEMICALS 17

4 *IN VIVO* REFERENCE DATA ON ACCURACY 18

4.1 TEST PROTOCOL *IN VIVO* REFERENCE TEST METHOD 18
4.2 ACCURACY 18
4.3 ORIGINAL RECORDS 19
4.4 QUALITY OF DATA 19
4.5 TOXICOLOGY 20
4.6 BACKGROUND ON ASSAY PERFORMANCE 20

5 TEST METHOD DATA AND RESULTS 21

5.1 TEST METHOD PROTOCOL 21
5.2 ACCURACY AND RELIABILITY 23
5.3 STATISTICS 29
5.4 TABULATED RESULTS 32
5.5 CODING OF DATA 34
5.6 CIRCUMSTANCES 34
5.7 OTHER DATA AVAILABLE 34

6 TEST METHOD ACCURACY 35

6.1 ACCURACY 35
6.2 CONCORDANCY TO *IN VIVO* REFERENCE METHOD 37
6.3 COMPARISON WITH REFERENCE METHODS 37
6.4 STRENGTH AND LIMITATIONS 37
6.5 DATA INTERPRETATION 38
6.6 COMPARISON TO OTHER METHODS 38

7 TEST METHOD RELIABILITY (REPEATABILITY/REPRODUCIBILITY) 39

7.1 SELECTION OF SUBSTANCES 39
7.2 RESULTS 39
7.3 HISTORICAL DATA 44
7.4 COMPARISON TO OTHER METHODS 44

8 TEST METHOD DATA QUALITY 45

8.1 CONFORMITY 45
8.2 AUDITS 45
8.3 DEVIATIONS 45
8.4 RAW DATA 45

9 OTHER SCIENTIFIC REPORTS AND REVIEWS 46

9.1 SUMMARY 46
9.2 DISCUSSION 47
9.3 RESULTS OF SIMILAR VALIDATED METHOD 48

10 ANIMAL WELFARE CONSIDERATIONS (REFINEMENT, REDUCTION, AND REPLACEMENT) 51

10.1 DIMINISH ANIMAL USE 51
10.2 CONTINUATION OF ANIMAL USE 51

11 PRACTICAL CONSIDERATIONS 53

11.1 TRANSFERABILITY 53
11.2 TRAINING 53
11.3 COST CONSIDERATIONS 54
11.4 TIME CONSIDERATIONS 54

12 REFERENCES 55

13 CATCH-UP VALIDATION: HUMAN WB/IL-1 IN VITRO PYROGEN TEST USING 96-WELLS PLATES. 57

13.1 RATIONALE 57
13.2 TEST METHOD PROTOCOL COMPONENTS 57
13.3 SUBSTANCES USED FOR VALIDATION 57
13.4 PRELIMINARY ESTIMATE OF THE TEST METHOD ACCURACY 57
13.5 TEST METHOD ACCURACY 61
13.6 TEST METHOD RELIABILITY (REPRODUCIBILITY) 65
13.7 CONCLUSION 65

14 SUPPORTING MATERIALS (APPENDICES) 66

14.1 STANDARD OPERATING PROCEDURE (SOP) OF THE PROPOSED METHOD 66
14.2 STANDARD OPERATING PROCEDURE (SOP) OF THE REFERENCE METHOD 66
14.3 PUBLICATIONS 66
14.4 ORIGINAL DATA 71
14.5 PERFORMANCE STANDARDS 72

Appendix A – Method protocol(s) and trial plan(s)

Appendix B – Hardcopies of relevant publications

Appendix C – List of abbreviations and definitions

1 Rationale for the Proposed Test Method

1.1 Introduction

1.1.1. *Describe the historical background for the proposed test method, from original concept to present. This should include the rationale for its development, an overview of prior development and validation activities, and, if applicable, the extent to which the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards.*

Pyrogens, a chemically heterogeneous group of hyperthermia- or fever-inducing compounds, derive from bacteria, viruses, fungi. Subjects react to such microbial products during an immune response by producing endogenous pyrogens such as prostaglandins and the pro-inflammatory cytokines interleukin-1 (IL-1), interleukin-6 (IL-6) and tumour necrosis factor-α (TNF-α) (Dinarello, 1999). Depending on the type and amount of pyrogen challenge and the sensitivity of an individual, even life-threatening shock-like conditions can be provoked. To assure quality and safety of any pharmaceutical product for parenteral application in humans, pyrogen testing is therefore imperative.

Depending on the medicinal product, one of two animal-based pyrogen tests is currently prescribed by the respective Pharmacopoeias, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET). For the rabbit pyrogen test, sterile test substances are injected intravenously to rabbits and any rise in body temperature is assessed. This *in vivo* test detects various pyrogens but not alone the fact that large numbers of animals are required to identify a few batches of pyrogen-containing samples argues against its use. In the past two decades, the declared intention to refine, reduce and replace animal testing, has lowered rabbit pyrogen testing by 80% by allowing to use the BET as an *in vitro* alternative pyrogen test for certain medicinal products (Cooper et al, 1971).

Bacterial endotoxin, comprising largely lipopolysaccharide (LPS) from the cell wall of Gram-negative bacteria that stimulates monocytes/macrophages via interaction with CD14 and toll-like receptor 4 (TLR4) (Beutler and Rietschel, 2003), is the pyrogen of major concern to the pharmaceutical industry due to its ubiquitous sources, its stability and its high pyrogenicity. With the BET, endotoxin is detected by its capacity to coagulate the amoebocyte lysate from the haemolymph of the American horseshoe crab, *Limulus polyphemus*, a principle recognised some 40 years ago (Levin and Bang, 1964). In the US, *Limulus* crabs are generally released into nature after drawing about 20% of their blood and therefore most of these animals survive. However, the procedure still causes mortality of about 30,000 horseshoe crabs per year, which adds to even more efficient threats of the horseshoe crab population like its use as bait for fisheries, habitat loss and pollution (http://www.horseshoecrab.org/).

As with the rabbit test the general problem of translation of the test results to the human fever reaction persists. Moreover, although being highly sensitive, the failure of the BET to detect non-endotoxin pyrogens as well as its susceptibility to interference by e.g. high protein or lipid levels of test substances or by glucans impedes full replacement of the rabbit pyrogen test. Hence, hundreds-of-thousands rabbits per year are still used for pyrogen testing.

A test system that combines the high sensitivity and *in vitro* performance of the BET test with the wide range of pyrogens detectable by the rabbit pyrogen test is therefore required in order to close the current testing gap for pyrogen and to avoid animal-based tests. With this intention and due to improved understanding of the human fever reaction (Dinarello, 1999), test systems based on *in vitro* activation of human monocytes were developed. First efforts date back about 20 years, when peripheral blood mononuclear cells (PBMC) were used to detect endotoxin by monitoring the release of pyrogenic cytokines (Duff and Atkins, 1982; Dinarello et al 1984). Meanwhile, a number of different test systems, using either whole blood, peripheral blood mononuclear cells (PBMCs) or the monocytoid cell lines MONO MAC 6 (MM6) or THP-1 as a source for human monocytes and various read-outs were established (Poole et al., 1988; Ziegler et al, 1988; Tsuchiya et al, 1980; Hartung and Wendel, 1996; Hartung et al, 2001; Poole et al, 2003). These test systems were validated with the aim of developing a tool for formal inclusion into Pharmacopoeias, an important basis for implementing novel alternative pyrogen tests for product-specific validation.

1.1.2 *Summarize and provide the results of any peer review conducted to date and summarize any ongoing or planned reviews.*
All of the five methods are currently under peer review of the ECVAM Scientific Advisory Committee.

1.1.3 *Clearly indicate any confidential information associated with the test method; however, the inclusion of confidential information is discouraged.*
This document does not contain any confidential information.

1.2 Regulatory rationale and applicability

1.2.1 *Describe the current regulatory testing requirement(s) for which the proposed test method is applicable.*
To assure quality and safety of pharmaceutical products for parenteral application in humans, pyrogen testing is imperative. Depending on the drug, one of two pyrogen tests is currently prescribed by the European Pharmacopoeia and other international guidelines, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET).

1.2.2 *Describe the intended regulatory use(s) (e.g., screen, substitute, replacement, or adjunct) of the proposed test method and how it will be used to substitute, replace, or complement any existing regulatory testing requirement(s).*
Dependent on the product and the presence of relevant clinical data on unexpected pyrogenicity of clinical lots, the proposed test method may be an alternative method for pyrogen testing, thus substituting the rabbit pyrogen test or the BET. In certain cases, the proposed test method may function as a supplementary test method to assess compliance to the licensing dossier.
In case the proposed test method is an alternative for pyrogenicity testing, a thorough cross-validation between the proposed test method and the original method for the specific medicinal product is warranted. In case the proposed test method is an adjunctive test to screen for (unexpected) pyrogenic lots, alert and alarm limits may be established based on consistency of production lots or (preferably) based on actual clinical data.

1.2.3 *Where applicable, discuss the similarities and differences in the endpoint measured in the proposed test method and the currently used in vivo reference test method and, if appropriate, between the proposed test method and a comparable validated test method with established performance standards.*
The current *in vivo* method (rabbit test), as described in the pharmacopoeia, and the proposed *in vitro* test method each determine very different end-points, though the biochemical origins of the response are similar.
The *in vivo* method more resembles a black box, and determines the total rise in body temperature (fever induction) of the animals subjected to the medicinal product, as a result of pyrogens (if any) present in the product.
The proposed test method WB/IL-1 is an *in vitro* model for the fever response mechanism. It determines the release of cytokines by monocytoid cells into the culture medium upon the interaction of pyrogens and specific receptors on the monocytoid cells. It is these cytokines that trigger the fever response *in vivo*.
Main differences between the *in vivo* and *in vitro* methods are that the latter is quantitative and uses cells of human origin, thus better reflecting the physiological situation.

1.2.4 *Describe how the proposed test method fits into the overall strategy of hazard or safety assessment. If a component of a tiered assessment process, indicate the weight that should be applied relative to other measures.*
The proposed test method WB/IL-1 may be applied for those medicinal products for which the rabbit test is the only or most reliable method for pyrogenicity testing, since a) the medicinal product is not compatible with the BET or b) the medicinal product contains pyrogens other than Gram-negative endotoxin.
Limit concentrations for pyrogens are established based on consistency lots or actual clinical data or, in the case of endotoxin the Endotoxin Limit Concentration (ELC) as defined for many medicinal products.

1.3 Scientific basis for the proposed test method

1.3.1 *Describe the purpose and mechanistic basis of the proposed test method.*
The proposed *in vitro* method is intended to determine the presence of pyrogens in medicinal products for parenteral use. The proposed test method is an *in vitro* model of the human fever response. It determines the release of cytokines upon the interaction of pyrogens and specific Toll-like receptors on the monocytoid cells (Beutler and Rietschel, 2003). These cytokines trigger the fever response *in vivo*.

1.3.2 *Describe what is known and not known about the similarities and differences of modes and mechanisms of action in the proposed test method as compared to the species of interest (e.g., humans for human health-related toxicity testing).*
An important feature of the proposed test method is that it is based upon the use of monocytoid cells of human origin. It therefore by definition resembles more closely the actual response of humans. The two other test methods make use of either crustaceans (BET) or rabbits, both species more or less distinct from the human species. The response of humans, horseshoe crabs and rabbits toward Gram-negative endotoxin has been

studied extensively and the methods appear equivalent for this particular pyrogen (Cooper et al 1971; Greisman and Hornick, 1969). However, there are documented cases of medicinal products and specified pyrogenic substances that yield false-positive or false-negative results in either test method. Since the proposed test method is based on human cells, it may therefore predict more accurately the pyrogenicity of such substances in humans.

1.3.3 *Describe the intended range of substances amenable to the proposed test method and/or the limits of the proposed test method according to chemical class or physicochemical factors.*

The proposed test method is intended for the assessment of pyrogens in all parenteral medicinal products for human use, chemical or biological and including raw materials, bulk ingredients and excipients. Use of the proposed test method in testing environmental samples or medicinal products is suggested and may be feasible, but substantiating data are as yet limited or absent.

2 Test Method Protocol Components

2.1 Overview of test method.

Provide an overview of how the proposed test method is conducted. If appropriate, this would include the extent to which the protocol for the proposed test method adheres to established performance standards.

A highly detailed protocol of the proposed test method (*Detailed protocol WB/IL-1: "Human whole blood pyrogen test"; electronic file name: SOP WB-IL-1*) is attached in Appendix A of this background review document (BRD). Appendix A also includes the amended protocol used in the formal validation study to determine the sensitivity and specificity of the test (section 3, table 3.3.1). However, it does only replace the previous version for testing of parenteral drugs described in table 3.3.1, and was included into Appendix A for completeness of information only *("Human Whole Blood Pyrogen Test - Standard Operating Procedure for the Validation Phase" marked with internal identifierSop-WBT-KNv02; electronic file name: SOP WB-IL-1 validation)*.

The WB/IL-1 test method is a two-part assay for the detection of pyrogenic contamination. The test protocol itself can be divided into the following two parts:

1. Incubation of the sample with (diluted) human blood.
2. An enzyme linked immunoassay (ELISA) for the measurement of IL-1β.

Ad 1.
Human whole blood from a single healthy volunteer is collected by venipuncture into heparinized tubes for blood sampling and used within 4 hours. Diluted human whole blood is incubated overnight (10-24 hours) together with saline and the sample of interest in sterile and pyrogen-free reaction tube. The supernatant is subsequently collected for further examination.

Ad 2.
Samples (supernatants of blood stimulation) are distributed into the wells of a microtiterplate which are coated with monoclonal antibodies specific for IL-1.
An enzyme-conjugated polyclonal antibody against IL-1β is added. During a 90-minute incubation, a sandwich complex consisting of two antibodies and the IL-1β is formed. Unbound material is removed by a wash step.
A chromogenic substrate (3,3´,5,5´ -tetramethylbenzidine, TMB) reactive with the enzyme label is added. Color development is terminated by adding a stop solution after 30 minutes. The resulting color, read at 450 nm, is directly related to the IL-1β concentration.

The WHO-LPS standard (code 94/580, E.coli O113:H10:K-), was used throughout the validation. This standard is identical to USP Reference Standard Endotoxin (EC6).
There are several possibilities to estimate the pyrogenic contamination of the preparations under test: 1) A quantitative estimation can be achieved by the construction of a dose-response curve for endotoxin standard (e.g. 5.0, 2.5, 1.0, 0.5 and 0.25 EU/ml) versus

Optical density (OD) value of the IL-1β ELISA. The contamination of the preparations is expressed in endotoxin–equivalent units. 2) A qualitative test can be achieved by the inclusion of an endotoxin threshold control (e.g. one fixed dilution of the standard curve) which allows for the classification in positive and negative samples (i.e. pyrogenic and non-pyrogenic samples). 3) A qualitative test can also be achieved by inclusion of an appropriate positive product control.
A detailed description of analysis methods used during the validation of the test method can be found in section 5 of the current BRD.

2.2 Rational for selected test components

Provide a detailed description and rationale, if appropriate, for the following aspects of the proposed test method:

2.2.1 Materials, equipment, and supplies needed.
The materials, equipment and supplies used for the WB/IL-1 test method are laboratory items, that will be already available in a routine QC laboratory. There is no need for sophisticated or dedicated laboratory equipment throughout the test.
For all steps in the procedure, excluding the ELISA procedure, the materials (e.g. tips, containers, solutions) which will be in close contact with samples and blood cells need to be sterile and pyrogen free. The materials, equipment and supplies are specified in the detailed protocol attached in Appendix A. It should be noted that equivalent devices may also be used and it is the user's responsibility to validate the equivalence.

Materials for part 1: Blood Incubation
Equipment
- Incubator or thermoblock (37°C ± 1°C)
- Adjustable 100 to 1000 µl (multi)pipettes
- Centrifuge (recommended)
- Vortex mixer

Consumables
- Heparinized tubes for blood sampling, e.g. Sarstedt S-MONOVETTE 7.5 ml, 15 IU/ml Li-Heparin
- Sarstedt multifly needle set, pyrogenfree, for S-Monovette
- 1.5 ml closable, pyrogen-free reaction tubes, e.g. from Eppendorf
- Reservoir for saline
- 12 ml or 15 ml tubes, e.g. from Greiner bio-one, for dilution of substances
- Sterile and pyrogen-free tips 100 µl and 1000 µl
- 10 ml and 2.5 ml pipets

Materials for part 2: ELISA procedure
Equipment
- Multichannel pipettor
- Microplate mixer
- Microplate washer
- Microplate reader capable of readings at 450 nm (optional reference filter in the range of 600-690 nm)

- A software package facilitating data generation, analysis, reporting, and quality control

Consumables

- Graduated cylinder and plastic storage container for Buffered Wash Solution
- Tip-Tubs for reagent aspiration with Multichannel pipettor
- The IL-1β-ELISA kit (commercially obtained), containing:
  - IL-1β antibody coated micoplates.
  - Enzyme labeled antibody. Horseradish peroxidase-labeled, affinity-purified, polyclonal(rabbit) anti-IL-1β antibodies.
  - Endotoxin control.
  - Saline
  - TMB/Substrate solution
  - Buffered Wash Solution Concentrate (saline solution, with surfactants and preservative)
  - Stop Solution (acidic solution)

2.2.2 Dose-selection procedures, including the need for any dose range-finding studies or acute toxicity data prior to conducting a study, if applicable.

For every kind of test compound the interference with human blood and the Il-1β ELISA kit is determined. For this purpose, a preliminary "dose finding" test is conducted to establish a suitable (interference free) dilution for every new test compound. For the validation study (as described in section 4 of this BRD), the tested products were diluted according to their known ELC, which was usually far beyond interfering concentrations. The ELCs of the tested products or drugs were calculated according to the European Pharmacopoeia.

2.2.3 Endpoint(s) measured.

The proposed test method is an *in vitro* model of the fever response mechanism. It determines the release of interleukin-1β (IL-1β) by monocytoid cells present in human blood. IL-1β is released into the culture medium upon the interaction of pyrogens and specific receptors on the monocytoid cells. The measured endpoint IL-1β is one of the cytokines that trigger the fever response *in vivo*.

2.2.4 Duration of exposure.

The human whole blood is exposed to possible pyrogenic components in samples overnight (10-24 hours) at 37°C. During the validation (described in section 4) the exposure of the cells was discontinued by centrifugation (2 minutes at 10,000 g) and collection of the clear supernatant. This supernatant, containing endogenous pyrogens released by the cells, is subsequently assayed in the IL-1β ELISA.

2.2.5 Known limits of use.

The WB/IL-1 method described in the method protocol is not a finalized test system for the testing of all medicinal products. The method may be applied only to preparations that have been validated with this method, i.e. shown not to interfere with the blood and the IL-1β readout system at a specified dilution of the preparation. A paragraph describing the interference testing is included in the method protocol (see Appendix A).

IL-1β readout system at a specified dilution of the preparation. A paragraph describing the interference testing is included in the method protocol (see Appendix A). However, at this moment there are no medicinal products known that cannot be tested with the method.

2.2.6 Nature of the response assessed.
The proposed test method is an *in vitro* model of the fever response mechanism. Upon the interaction of exogenous pyrogens and specific receptors on the monocytoid cells endogenous pyrogens (e.g. interleukins, TNF-α and prostaglandins) are produced. In the body the fever response is triggered by these endogenous pyrogens. Immunoreactive IL-1β, the measured endpoint for the current method, is one of these endogenous pyrogens.

2.2.7 *Appropriate vehicle, positive, and negative controls and the basis for their selection.*
Throughout the development and validation phase the test compounds are diluted in 0.9% (w/v) clinical saline. This 0.9% clinical saline is considered an appropriate vehicle as no interference with active substances of a drug is to be expected.
In addition the test includes several controls.
A negative control: 0.9% clinical saline (sodium chloride)
A positive control: WHO-LPS 94/580, 0.5 EU/ml in clinical saline.
A negative product control: clean, released batch for each drug.
A positive product control: test item spike with WHO-LPS (code 94/580) at 0.5 EU/ml
The positive and negative controls are the same in every assay and are needed to establish the sensitivity of the test system. In addition, a product-based set of controls is used to reveal product-related interference.

2.2.8 Acceptable range of vehicle, positive and negative control responses and the basis for the acceptable ranges.
A WB/IL-1 assay is considered acceptable for further analysis if the mean OD of the positive control (0.5 EU/ml) exhibits an OD that is greater than 1.6 times the mean OD over the negative control (0.9% clinical saline). Moreover the response to different concentrations of the positive control should show a dose response relationship. To be able to quantify the responses to the positive control this should be well within the maximum response that can be measured with the test system.

As regards the substances to be tested, for products with an established ELC (specified in EU/ml), the product is diluted to its maximum valid dilution (MVD). The negative product control should be negative at the MVD. The response to the positive product control should be between 50% and 200% of the response to the positive control, indicating a possible pyrogenicity can be detected using these conditions.

2.2.9 Nature of the data to be collected and the methods used for data collection.
The raw data collected are the read-outs (absorbance) of the IL-1β ELISA, measured by an automated laboratory ELISA-plate reader. The wavelength is dependent on the chromogenic substrate applied, but when using TMB, the ELISA-plate is read at a

2.2.10 Type of media in which data are stored.
Data are stored in electronic files (windows98 compatible software) and as hard copy.

2.2.11 Measures of variability.
As part of the development of the WB/IL-1 test method the intralaboratory repeatability was assessed by independent and identical replicated measurement of the different concentrations of WHO-LPS. Furthermore, the limit of detection and its dependence from known but uncontrollable variables such us operator and passage of the cell line were investigated. These variables and the inherent variation of biological systems make up to the total variation of the method.

2.2.12 Statistical or nonstatistical methods used to analyze the resulting data, including methods to analyze for a dose-response relationship. Justify and describe the method(s) employed.
All experiments are run with four replicates of the test compound with blood from one donor on one plate. A standard curve in quadruplicate, using the International Standard for endotoxin (calibrated in EU) is included, ranging from 0.25 EU/ml up to 2.5 EU/ml. Outliers are rejected only after checking according to the Grubbs test, and applied to identify and eliminate aberrant data. Next, the negative and the respective positive control are compared to ensure a suitable limit of detection, which should be >0.25 EU/ml.
The mean OD of the 0.5 EU/ml endotoxin control exhibits an OD that is greater than 1.6x the OD of the negative saline control.

2.2.13 Decision criteria and the basis for the prediction model used to classify a test chemical (e.g., positive, negative, or equivocal), as appropriate.
A prediction model (PM) was developed in order to classify substances as “pyrogenic for humans” or “non-pyrogenic for humans”. To be able to define a dichotome result in the alternative pyrogen test, a threshold pyrogen value of 0.5 EU/ml was chosen. This threshold value was based on historical data with rabbits (described in section 4.1). The suitability of the PM was assessed by testing substances which were artificially contaminated with endotoxin (substances are described in section 3.2 and 3.3).
The statistical approach, including quality criteria, is detailed in section 5.3

2.2.14 Information and data that will be included in the study report and availability of standard forms for data collection and submission.
Raw data were collected using a standard form. These were submitted to the quality department of ECVAM.

2.3 Basis for selection of this test method

Explain the basis for selection of the test method system. If an animal model is being used, this should include the rationale for selecting the species, strain or stock, sex, acceptable age range, diet, and other applicable parameters.

In view of the shortcomings of the rabbit pyrogen test and the BET, *in vitro* pyrogen tests that utilize the exquisite sensitivity to exogenous pyrogen of monocytoid cells have been proposed. In such tests, products are incubated with human cell and the conditioned media assayed for pyrogenic cytokines (Duff & Atkins, 1982; Dinarello et al., 1984; Poole et al., 1988; Poole et al, 1989; Hansen and Christensen, 1990; Taktak et al., 1991; Bleeker et al., 1994).

The human whole blood assay was developed as a real *in vitro* alternative to the rabbit pyrogen test. The basic idea was to mimic the fever reaction in humans. In general, the detection of exogenous pyrogens (e.g. endotoxin) by blood cells causes them to release endogenous pyrogens like IL-1β, IL-6 and TNF α. These cytokines affect the thermal regulation centre in the brain and increase the body temperature by changing its set point. In the past several test methods have been developed that use the sensitivity of human peripheral blood monocytes to exogenous pyrogens. In an attempt to increase the sensitivity of these tests the monocytes/leukocytes were isolated from whole blood. In addition, various cell lines, which retain monocytoid characteristics, including the capacity to synthesize and secrete pyrogenic cytokines, have been studied.

However, the isolation of monocytes/leukocytes from whole blood as well as the maintenance of a cell-line is labour–intensive and time–consuming, technically sophisticated and requires expensive reagents. It is clear that using whole blood implies considerably simplified handling and that costs are limited. In an early stage of development of the assay, interleukin-1β was most promising as the endogenous pyrogen used as the readout. In addition, a standardised version of the test in form of an interleukin-1β kit is commercially available.

2.4 Proprietary components

If the test method employs proprietary components, describe what procedures are used to ensure their integrity (in terms of reliability and accuracy) from "lot-to-lot" and over time. Also describe procedures that the user may employ to verify the integrity of the proprietary components.

T. Hartung and A. Wendel are named as inventors in Patent Number US 5,891,728 , Apr 6, 1999: 'Test for determining pyrogenic effect of a material'.

It is stated in the method protocol that components supplied in the ELISA kit are not interchangeable with other lots of the same components. Including the appropriate positive and negative controls in each run ensures the reliability and accuracy of the WB/IL-1 test method. As a positive control a specified amount of the Endotoxin Standard is used. The assay should be considered acceptable only if the following criteria are met: The mean absorbance of the 0.5 EU/ml endotoxin control exhibits a value is greater than 1.6x the mean absorbance of the negative saline control. Requirements are set for variability of replicates within an assay.

In addition the response to the negative control should be well below limit of detection.

2.5 Replicates

Describe the basis for the number of replicate and repeat experiments; provide the rationale if experiments are not replicated or repeated.

All experiments are run with four replicates of the test compound on one plate. Outliers are rejected only after checking according to the Grubbs test ($p>0.05$). Four replicates is considered the minimal amount for the Grubbs test.

During a prevalidation phase, the intralaboratory reproducibility as well as the interlaboratory reproducibility of the WB/IL-1 test method was established by applying repeated experiments (see section 7). As the test method reliability (repeatability /reproducibility) was shown to be satisfactory, it was feasible to establish the accuracy using pharmaceutical substances (detailed in table 3.3.1) by one test performed by three participating laboratories (see section 6).

2.6 Modifications applied after validation

Discuss the basis for any modifications to the proposed test method protocol that were made based on results from validation studies.

The highly detailed method protocol (Appendix A) also allows for omission of the centrifugation step of the blood/sample mixture. However, in order to reduce the variability of the assay to a minimum, the separation of blood and supernatant was obligatory for the final validation study (Appendix A; see also 2.1).

The test can easily be adjusted to a quantitative assay as described in the detailed method protocol. However, the assay has now been validated as a qualitative assay, by means of the PM.

2.7 Differences with similar test methods

If applicable, discuss any differences between the protocol for the proposed test method and that for a comparable validated test method with established performance standards.

Not applicable.

3 Substances Used for Validation

3.1 Selection of substances used

Describe the rationale for the chemicals or products selected for use in the validation process. Include information on the suitability of the substances selected for testing, indicating any chemicals that were found to be unsuitable.

Selected test items were medicinal products available on the market. Released clinical batches were considered clean, i.e. containing no detectable pyrogens. To test the specificity, sensitivity and the reproducibility of the proposed test method, the products were spiked with pyrogen. For the present studies endotoxin (LPS) was selected as the model pyrogen, since it is well defined, standardized and readily available.

For the sensitivity and specificity the test items were assessed at their MVD. The MVD is the quotient of the ELC and the detection limit. The European Pharmacopoeia prescribes for various types of parenterals the amount of endotoxin that is maximally allowed in a medicinal product, i.e. the ELC, taking into consideration the dose, the route of administration and the dosing regimen of the product.

The aim of the study was to discriminate between negative and positive samples. Based on the selected pyrogen threshold value (see 4.1) and the PM applied, positive samples were defined as containing 0.5 EU/ml or more. Hence, to determine the MVD, the value of 0.5 EU/ml was defined as the detection limit.

Test items were assessed as such (negative product control), spiked with endotoxin at 0.5 IU/ml (positive product control) and after spiking with endotoxin at 5 levels (blinded samples). In addition a negative control (saline) and positive control (0.5 EU/ml in saline) were included to establish assay validity.

For reproducibility, the test items were tested, at a predefined dilution above the MVD, independently in 3 different laboratories, 3 times each. Based on the selected pyrogen threshold value (see 4.1) and the PM applied, positive samples were defined as containing 0.5 EU/ml or more. The test items were tested after spiking with endotoxin at four levels. For no other reasons but practical ones, i.e. availability of test materials, different test items were selected for this part of the validation study.

It was determined earlier whether candidate test items interfered with the outcome of the proposed test method. Interference was considered when the response of endotoxin in the diluted test item was below 50% or above 200% of the response of endotoxin in saline (spike-recovery). It was shown that none of the test items interfered with the assay at the selected dilutions (data not shown).

3.2 Number of substances

Discuss the rationale for the number of substances that were tested.

A total of 13 test items were selected for the validation study (see 3.3): 10 test items for determining sensitivity and specificity (table 3.3.1), 3 different test items for determining reproducibility (table 3.3.2). Test items and their spikes were appropriately blinded by ECVAM before distribution to the participating testing facilities.

For sensitivity and specificity, each test item was tested after spiking at its individual MVD. Hence they each came with their own specific set of 5 endotoxin spike solutions: 0.0, 0.25, 0.5, 0.5 and 1.0 EU/ml. Simple logistics limited the amount of test items for this part of the validation study to 10. Since test items were assessed with 5 different endotoxin levels at 3 independent test facilities, this yielded a total of 150 data points, biometrically considered to be sufficient for further analysis.

For reproducibility each test item was spiked at 4 different levels (0.0, 0.0, 0.5 and 1.0 EU/ml) and tested at specified dilutions, 3 times at 3 laboratories.

3.3 Description of substances used

Table 3.3.1: Test items (parenteral drugs) used for determining sensitivity and specificity

Drug	code	Source	Agent	Indication	MVD (-fold)
Glucose 5% (w/v)	GL	Eifel	Glucose	nutrition	70
Ethanol 13% (w/w)	ET	B.Braun	Ethanol	diluent	35
MCP®	ME	Hexal	Metoclopramid	antiemetic	350
Orasthin®	OR	Aventis	Oxytocin	initiation of delivery	700
Binotal®	BI	Aventis	Ampicillin	antibiotic	140
Fenistil®	FE	Novartis	Dimetindenmaleat	antiallergic	175
Sostril®	SO	GlaxoSmithKline	Ranitidine	antiacidic	140
Beloc®	BE	Astra Zeneca	Metoprolol tartrate	heart dysfunction	140
Drug A*	LO	-	0.9% NaCl	-	35
Drug B*	MO	-	0.9% NaCl	-	70

*Drugs A and B were included as saline controls using notional ELCs.
Negative control: 0.9% clinical stock saline solution.
Positive control: WHO-LPS 94/580 (*E. coli* 0113:H10:K-), 0.5 EU/ml in clinical stock saline.

Table 3.3.2: Test items (parenteral drugs) used for determining reproducibility.

Drug	Source	Agent	Indication
Gelafundin®	Braun melsungen	Gelatin	Transfusion
Jonosteril ®	Fresenius	Electrolytes	Infusion
Haemate ®	Aventis	Factor VIII	Hemophilia

Negative control: 0.9% clinical stock saline solution.
Positive control: WHO-LPS 94/580 (*E. coli* 0113:H10:K-), 0.5 EU/ml in clinical stock saline.

BRD: WB/IL-1 March, 2006

3.4 Sample coding procedure

Describe the coding procedures used in the validation studies.

All test items are registered medicinal products and were obtained from a pharmaceutical supplier. Test items and endotoxin spiking samples were prepared, blinded where appropriate and coded under GLP by personnel from ECVAM, Italy. These were then taken over by the Paul-Ehrlich Institute, Germany, for allocation and shipment to each of the appropriate test facilities participating in the study.
For the sensitivity and specificity part of this study, test items and their respective spikes (5 per test item) were all blinded. For reproducibility testing, only the spikes (4) were blinded, the test items were not.

3.5 Recommended reference chemicals

For proposed test methods that are mechanistically and functionally similar to a validated test method with established performance standards, discuss the extent to which the recommended reference chemicals were tested in the proposed test method. In situations where a listed reference chemical was unavailable, the criteria used to select a replacement chemical should be described. To the extent possible, when compared to the original reference chemical, the replacement chemical should be from the same chemical/product class and produce similar effects in the in vivo reference test method. In addition, if applicable, the replacement chemical should have been tested in the mechanistically and functionally similar validated test method. If applicable, the rationale for adding additional chemicals and the adequacy of data from the in vivo reference test method or the species of interest should be provided.

The reference pyrogen material used was the international endotoxin standard WHO-LPS 94/580 (*E. coli* 0113:H10:K-). Where appropriate, the material was diluted in clinical saline solution (0.9%(w/v) sodium chloride). The saline was also used as negative control (blank).

4 *In vivo* Reference Data on Accuracy

4.1 Test protocol *in vivo* reference test method.

Provide a clear description of the protocol(s) used to generate data from the in vivo reference test method. If a specific guideline has been followed, it should be provided. Any deviations should be indicated, including the rationale for the deviation.

For ethical reasons, no rabbit pyrogen tests were performed for this study. However, Dr. U. Lüderitz-Püchel, Paul-Ehrlich Institute, Germany, kindly provided historical data, accumulated over several years, from 171 rabbits (Chinchilla Bastards). The respective Pharmacopoeia's do not prescribe a rabbit strain for the *in vivo* pyrogen test, but Chinchilla rabbits are reported as a relatively sensitive strain for pyrogen testing.

The rabbits were injected with endotoxin and their rise in body temperature over the next 180 minutes was recorded (figure 4.1.1). From these data it was established that 50% of the rabbits got fever when treated with endotoxin at 5 EU/kg (Hoffmann et al, 2005a). Fever in rabbits is defined as a rise in body temperature over 0.55°C. On the basis of these historical animal data and corrected for the maximal volume allowed in rabbits, i.e. 10 ml per animal, a pyrogen threshold value of 0.5 EU/ml was defined for the PM in the proposed test method.

4.2 Accuracy

Provide the in vivo reference test method data used to assess the accuracy of the proposed test method. Individual human and/or animal reference test data, if available, should be provided. Provide the source of the reference data, including the literature citation for published data, or the laboratory study director and year generated for unpublished data.

As mentioned, no animal studies were done for ethical reasons. However, a theoretical assumption on specificity and sensitivity can be made (Hoffmann et al, 2005a). Taking the prevalence of the different final concentrations of LPS in the 5 spikes into account (1.0 EU/ml: 20%; 0.5 EU/ml: 40%; 0.25 EU/ml: 20% and 0.0 EU/ml: 20%) and calculating probabilities of misclassification using the chosen study design and defined threshold of pyrogenicity, i.e. 0.5 EU/ml, the theoretical sensitivity of the rabbit pyrogen test is 57.9% and the theoretical specificity of the rabbit pyrogen test is 88.3%.

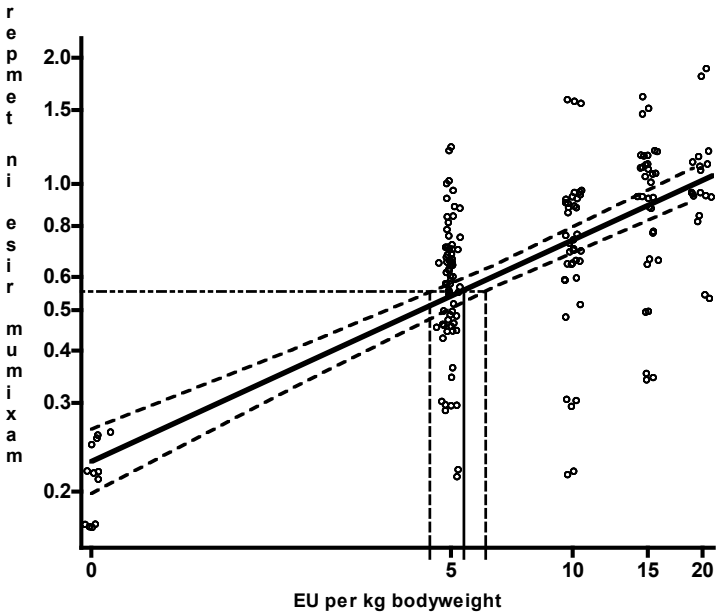


Figure 4.1.1 Dose-temperature of standard endotoxin applied to Chinchilla Bastards (n=171). Rabbits were treated with 1 ml saline containing 0, 5, 10, 15 and 20 EU of *E. coli* LPS (WHO-LPS 94/580 (E.coli O113:H10:K)) and their body temperature was measured over 180 min. Linear regression analysis was performed after logarithmic transformation of the data. Data are shown as dots to which a jitter-effect was applied in order to be able to distinguish congruent data. The full line depicts the linear regression whereas the dashed lines represent the 95%-confidence bounds. Furthermore, a horizontal line for a 0.55°C raise of temperature is added which is often defined as the rabbit threshold for fever. At the interception point of this line and the regression line 50% of the rabbits are to be expected to develop fever.

4.3 Original records

If not included in the submission, indicate if original records are available for the in vivo reference test method data.

The recognition of pyrogenic substances as bacterial by-products and the identification of a variety of pyrogenic agents enabled the development of a proper test to demonstrate non-pyrogenicity of the pharmaceutical product. As early as the 1920s, studies were done to select the most appropriate animal model. Results indicated that most mammals had a pyrogenic response, but only a few, including rabbits, dogs, cats, monkeys and horses showed a response similar to that in humans. For practical reasons, other species but rabbits and dogs were considered not practical. In 1942, Co Tui & Schrift described that rabbits are less thermo-stable as compared to dogs. Hence, rabbits are more suited for the purpose of testing for the absence of pyrogens, since a negative result is more significant.

4.4 Quality of data

Indicate the quality of the in vivo reference test method data, including the extent of GLP compliance and any use of coded chemicals.

Documented procedures were employed that were GLP-concordant. These were quality assured by quality assurance officers from ECVAM.

4.5 Toxicology

Discuss the availability and use of relevant toxicity information from the species of interest (e.g., human studies and reported toxicity from accidental or occupational exposure for human health-related toxicity testing).

Over time, a number of studies were done to correlate the rabbit test to pyrogenic reactions in humans. A conclusive study by Greisman and Hornick, published in 1969, who compared three purified endotoxin preparations (*Salmonella typhosa*, *E. Coli* and *Pseudomonas*) in New Zealand rabbits and in male volunteers, showed that the induction of a threshold pyrogenic response, on a weight basis, was similar to rabbit and man. At higher doses, rabbits respond less severe as compared to man.

4.6 Background on assay performance

Discuss what is known or not known about the accuracy and reliability of the in vivo reference test method.

As mentioned, no animal studies were done for ethical reasons. However, a theoretical assumption on specificity and sensitivity can be made (Hoffman et al, 2005a) Taking the prevalence of the different final concentrations of LPS in the 5 spikes into account (1.0 EU/ml: 20%; 0.5 EU/ml: 40%; 0.25 EU/ml: 20% and 0.0 EU/ml: 20%) and calculating probabilities of misclassification using the chosen study design and defined threshold of pyrogenicity, i.e. 0.5 EU/ml, the theoretical sensitivity of the rabbit pyrogen test is 57.9% and the theoretical specificity of the rabbit pyrogen test is 88.3%.

5 Test Method Data and Results

5.1 Test method protocol

Describe the proposed test method protocol used to generate each submitted set of data. Any differences from the proposed test method protocol should be described, and a rationale or explanation for the difference provided. Any protocol modifications made during the development process and their impact should be clearly stated for each data set.

The detailed method protocol for the WB/IL-1 test is provided in the Appendix A of this BRD. It includes the precise step-by-step description of the test method, including the listing of all the necessary reagents and laboratory procedures for generating data. For two steps during validation a part of the protocol was adapted to contain a detailed description of the dilution of the samples and the spiking with WHO-LPS. The relevant part of the protocol is detailed in this section as well. Both protocols (see also 2.1) are attached in the Appendix A. The validity criteria and the detailed statistical analysis described in section 5.3 of this BRD were applied to analyse the data produced during validation.

To assess the reliability of the test method a series of experiments were conducted in the developing laboratory (DL). As a start, only blanks (saline, 0 EU/ml) and spikes of WHO-LPS in saline were tested. These experiments are summarised in table 5.1.1.

Table 5.1.1 : summary of experiments with WHO-LPS in saline

Experiment	Spikes (EU/ml) in saline	n (per spike)	Repetitions of experiment	N
1A	0; 0.25; 0.5	20	1	60
1B	0; 0.063; 0.125; 0.25; 0;5	12	1	60
2A	0; 0.25; 0.5	20	3	180
2B	0; 0.25; 0.5	20	3	180

The collected data were used to answer questions regarding the nature of the distribution, the variance and its behaviour over the range of response in replicated measurements under identical conditions. In addition, intralaboratory reproducibility was assessed by the maintenance of minimum criteria, e.g. a detection limit below an a-priori chosen positive control or a dose-dependent standard curve (table 5.1.1, experiment 1b). With the data of this experiment an assessment of the limit of detection of the test can be done by calculating the smallest spike, which can be discriminated from the blank. Intralaboratory reproducibility was assessed by the maintenance of minimum criteria, e.g. a detection limit below an a-priori chosen positive control of a dose dependent standard curve.

The WB/IL-1 method was transferred from the DL to two other laboratories (denoted as naive laboratory 1 [NL1] and naive laboratory 2 [NL2]). A large-scale dose response experiment was performed by all three laboratories. For this study 6 or 7 concentrations were tested in a dose response curve (typically 0; 0.125; 0.25; 0.5; 1; 2; 4 EU/ml, at least 8 replicates) and all laboratories had to meet the validity criteria as laid down in the method protocol before the studies with medicinal substances were conducted.

The **(intra- and interlaboratory) reproducibility** was assessed by testing 3 different medicinal substances, Gelafundin, Jonosteril and Haemate (described in table 3.3.2, section 3.3.). Test items and their spikes were appropriately blinded. Test items were tested, at a predefined dilution above the MVD, independently in 3 laboratories, 3 times each. Test items were tested after spiking with WHO-LPS at four different levels, the spikes were blinded and coded by QA ECVAM. In addition a negative control (saline) and positive control (0.5 EU/ml) in saline were included to establish assay validity. Although this part of the study was designed for assessment of reproducibility, a preliminary estimate of the accuracy could be derived from the data. Applying the PM to the results and evaluating the concordance in a two-by-two contingency table assessed accuracy.

To assess **accuracy** of the proposed test method 10 substances (listed in table 3.3.1), were spiked with five different concentrations of the WHO-LPS (one of which is negative). To permit the application of the chosen PM, each drug was diluted to its individual MVD, which has been calculated using the ELCto that drug (listed in table 3.3.1.) Each substance had their own specific set of endotoxin spike solution, ensuring that after spiking the undiluted drug, it contained 0.0; 0.25; 0.5; 0.5 and 1.0 EU/ml respectively when tested at its MVD. A detailed description of the sample preparation was supplied to the three independent laboratories (shown in table 5.1.2). To put more weight to the result of this part of the validation, the spikes were blinded and coded by QA ECVAM. The raw data of the WB/IL-1 assay are shown in paragraph 5.2 Accuracy was assessed by applying the PM to the results and evaluating the concordance in a two by two table. As intralaboratory reproducibility was (successfully) shown in previous experiments, only interlaboratory reproducibility was assessed in this phase.

Table 5.1.2: Sample preparation for the testing of 10 substances spiked with 5 different concentrations of WHO-LPS.

unblinded			*blinded*				
dilution of drug up to **MVD** ↓			spiking of **undiluted** drug: 0.5 ml each				
diluted drug	**NPC**	**PPC**	+ 23.3 µl of Spike 1	+ 23.3 µl of Spike 2	+ 23.3 µl of Spike 3	+ 23.3 µl of Spike 4	+ 23.3 µl of Spike 5
0.5 ml	+ 25 µl saline	+ 25 µl PPC-LPS-spike *					
		(final conc. = 50 pg/ml)	**dilution** to **MVD** ↓				
	test	test	test	test	test	test	test

** PPC-LPS-spike contains 1050 pg/ml = 21fold 50 pg/ml*

NPC = Negative Product Control, **PPC** = Positive Product Control, **MVD** = Maximal Valid Dilution

BRD: WB/IL-1 March, 2006

5.2 Accuracy and reliability

Provide all data obtained to evaluate the accuracy and reliability of the proposed test method. This should include copies of original data from individual animals and/or individual samples, as well as derived data. The laboratory's summary judgment regarding the outcome of each test should be provided. The submission should include data (and explanations) from all studies, whether successful or not.
See figures 5.2.1, 5.2.2, 5.2.3, 5.2.4, 5.2.5 (A, B and C), 5.2.6 and 5.2.7 (A and B).

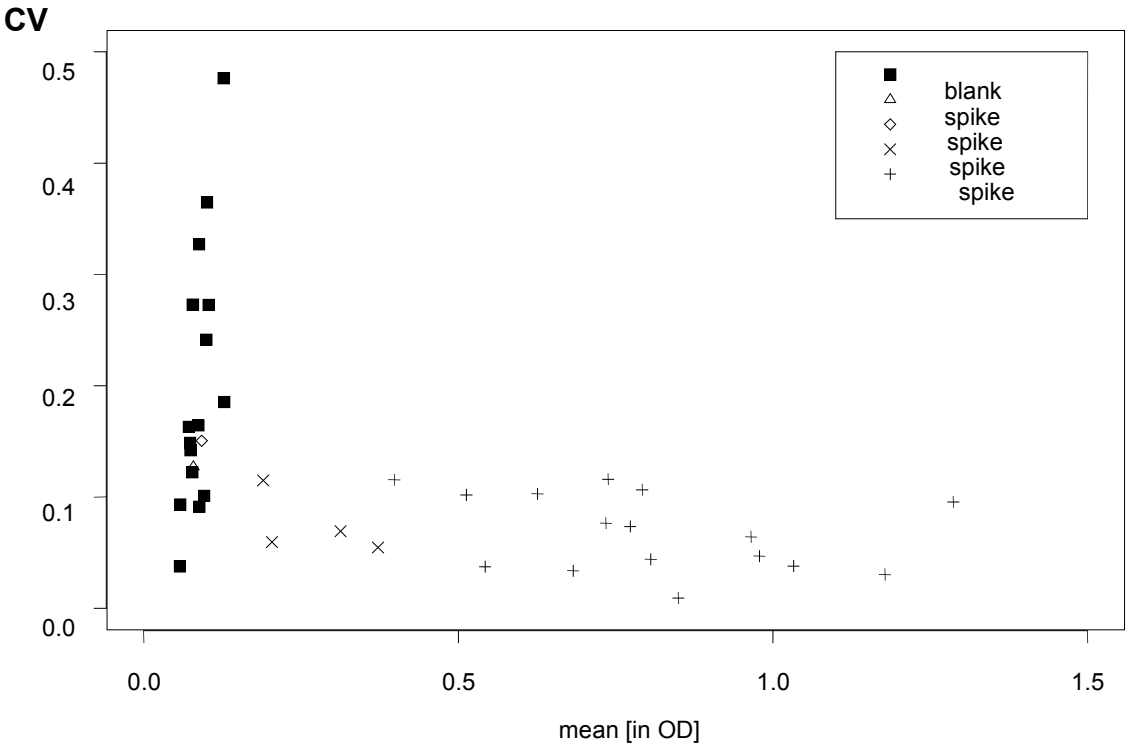


Figure. 5.2.1: Coefficient of variation (CV) of WHO-LPS spikes (4 replicates) relative to the mean OD (readout of the IL-1 ELISA).

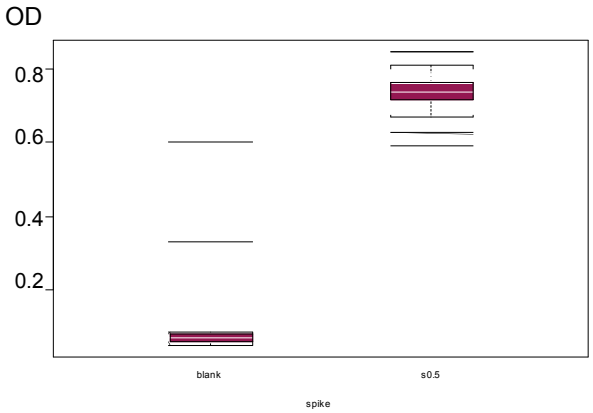


Figure. 5.2.2: Boxplots with OD values of 20- replicates (left) or 12 replicates (right) of WHO-LPS spikes in saline at various concentrations. (readout of the IL-1 ELISA).

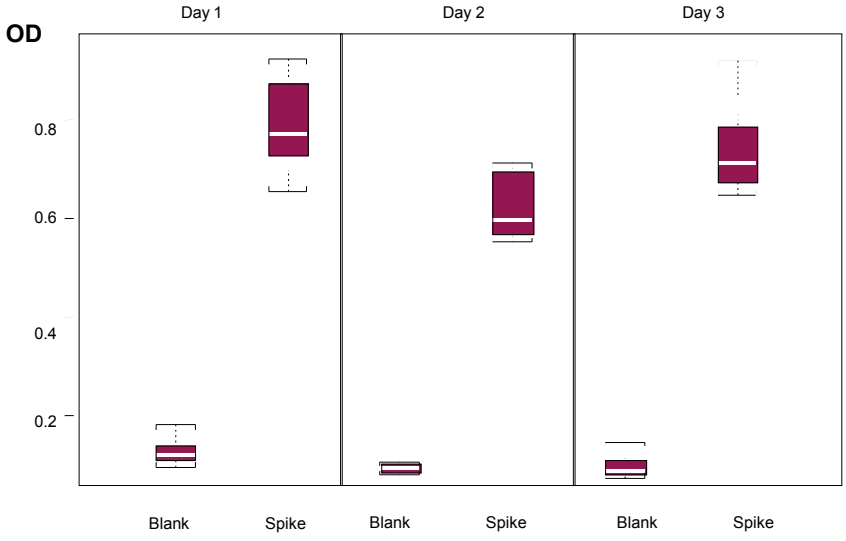


Figure. 5.2.3: Boxplots OD values of the response of 3 different blood donations from one healthy volunteer on consecutive days with WHO-LPS (IU/ml) in saline at 0.0 IU/ml (Blank) or 0.5 IU/ml (Spike) (readout of the IL-1 ELISA).

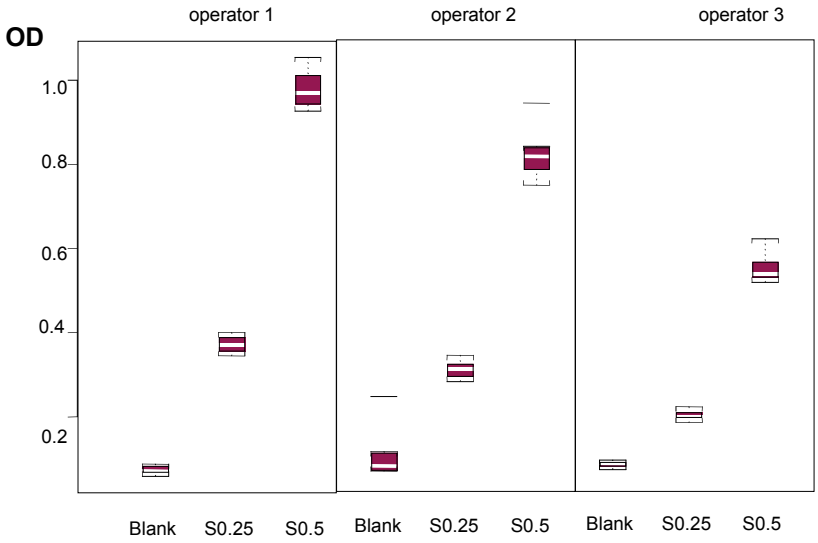


Figure. 5.2.4: Boxplots OD values of WHO-LPS (IU/ml) in saline at 0.0 (Blank) , 0.25 IU/ml (S0.25) or 0.5 IU/ml. (S0.5) with 3 different operators (readout of the IL-1 ELISA).

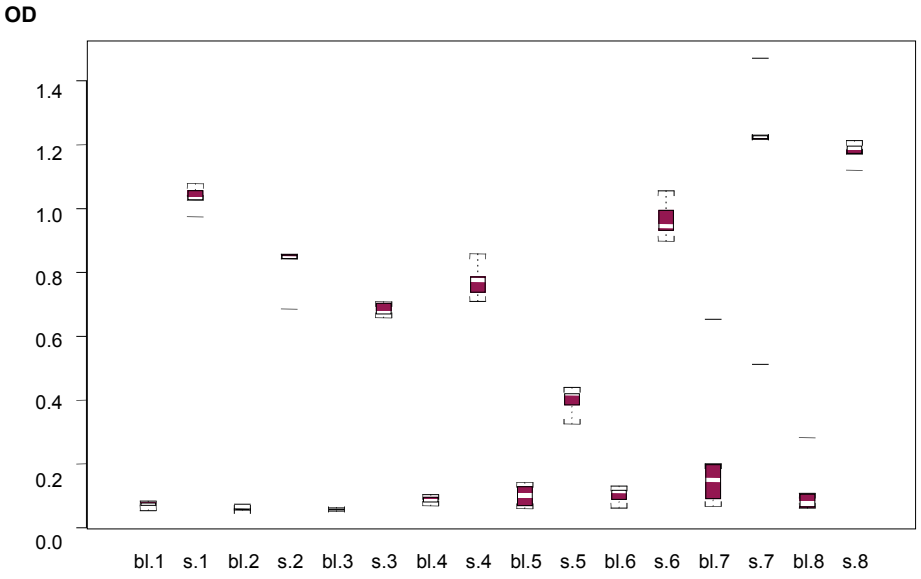


Figure. 5.2.5: Boxplots OD values of the response of 8 individual donors to WHO-LPS (IU/ml) in saline at 0.0 IU/ml (bl.*x*) or 0.5 IU/ml. (S.*x*) (readout of the IL-1 ELISA).

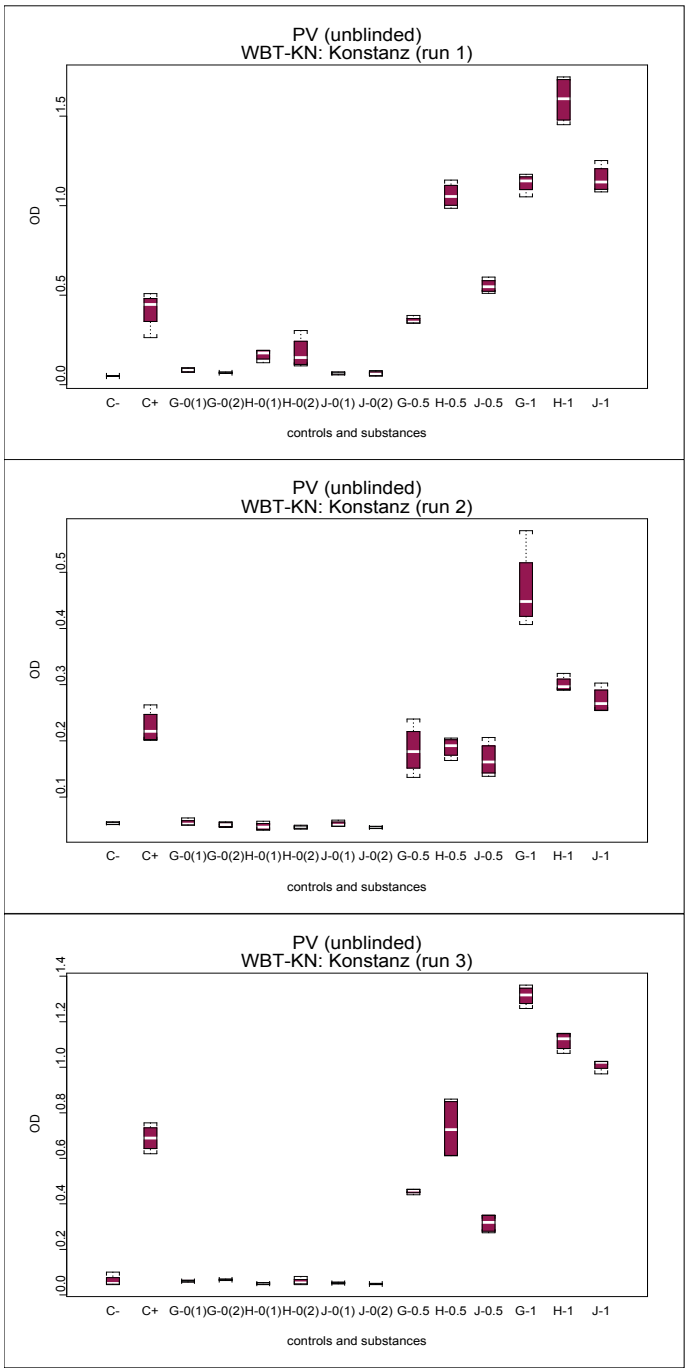


Figure. 5.2.6 A: Three different drugs were spiked (blinded) with WHO-LPS at 0.0, 0.0, 0.5 and 1.0 IU/ml, respectively. Experiment was run 3 time independently at three different laboratories. Here the results of the Konstanz laboratory (readout of the IL-1 ELISA).
G = Gelafundin; J = Jonestreril; H = Heamate.
C- = negative control (saline); C+ = positive control (0.5 IU/ml in saline).

BRD: WB/IL-1 March, 2006

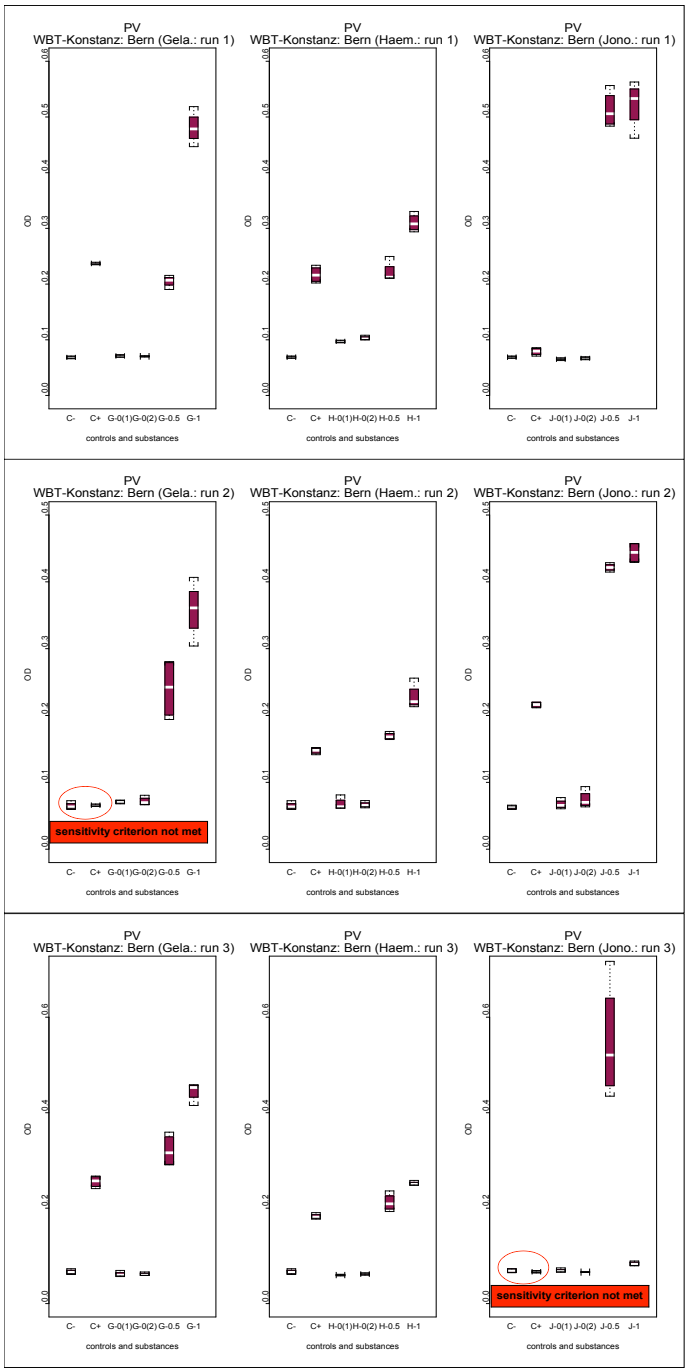


Figure. 5.2.6 B: Three different drugs were spiked (blinded) with WHO-LPS at 0.0, 0.0, 0.5 and 1.0 IU/ml, respectively. Experiment was run 3 time independently at three different laboratories. Here the results of the Bern laboratory (readout of the IL-1 ELISA). G = Gelafundin; J = Jonestreril; H = Heamate.
C- = negative control (saline); C+ = positive control (0.5 IU/ml in saline).

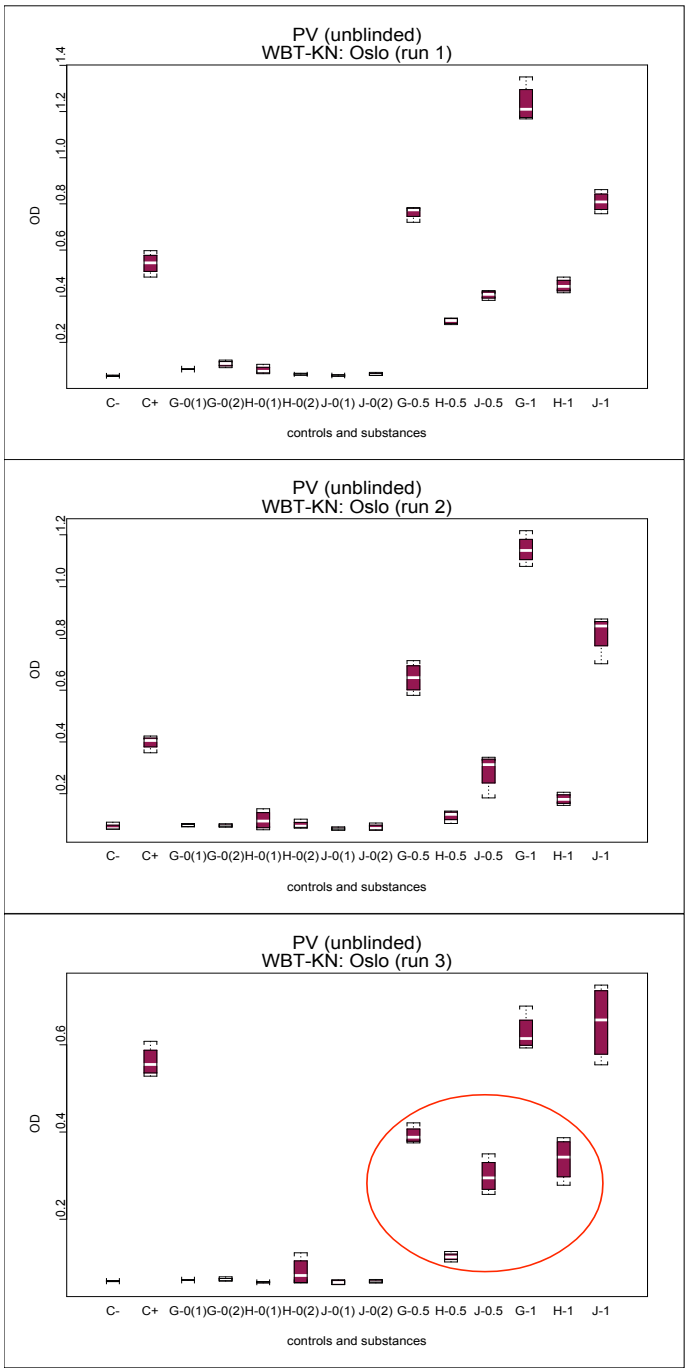


Figure. 5.2.6 C: Three different drugs were spiked (blinded) with WHO-LPS at 0.0, 0.0, 0.5 and 1.0 IU/ml, respectively. Experiment was run 3 time independently at three different laboratories. Here the results of the Oslo laboratory (readout of the IL-1 ELISA).
G = Gelafundin; J = Jonestreril; H = Heamate.
C- = negative control (saline); C+ = positive control (0.5 IU/ml in saline).

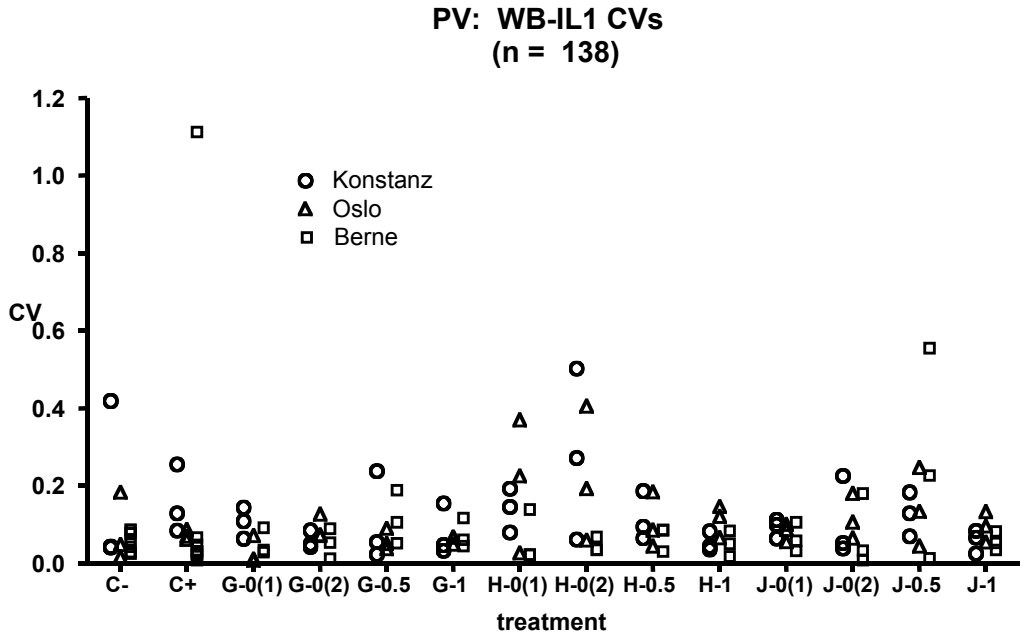


Figure. 5.2.7: Coefficient of variation (CV) of different WHO-LPS spikes (0.0, 0.0, 0.5 and 1.0 IU/ml, respectively).from the experiments as shown in fig. 5.2.6 A-C.
G = Gelafundin; J = Jonestreril; H = Heamate.
NC = negative controle (saline); PC is positive conrole (0.5 IU/ml in saline).

5.3 Statistics

Describe the statistical approach used to evaluate the data resulting from studies conducted with the proposed test method.

A generally applicable analytical procedure was employed. This procedure includes a universal PM as well as quality criteria. First, a two-step procedure consisting of a variance-criterion and an outlier-test was performed. For this, the Dixon's test (Barnett and Lewis, 1984), which is USP approved, was chosen with the significance level of α=0.01 and applied to identify and eliminate aberrant data.

Next, the negative and the respective positive control are compared to ensure a suitable limit of detection. For this, a one-sided t-test with a significance level of α=0.01 is applied to the ln-transformed data to ensure that the response to the positive control is significantly larger than that of the respective negative control.

Finally, the samples are classified as either negative or positive by the outcome of a one-sided version of the t-test, which is based on the assigned pyrogen threshold value. The final results will be given in 2 x 2 contingency tables (table 5.3.1). These tables allow for estimation of accuracy (sensitivity and specificity) and reproducibility of the proposed test method.

Table 5.3.1: 2x2 contingency table.

		pre-defined class ("truth")		Σ
		1	***0***	
Classification by test system and PM	***1***	***a***	***b***	$a+b=n_{.1}$
	0	***c***	***d***	$c+d=n_{.0}$
Σ		$a+c=n_{1.}$	$b+d=n_{0.}$	***n***

Accuracy:

The most important statistical tool to determine accuracy (specificity and sensitivity) is the so-called PM (Hothorn, 1995). In general, it is a statistical model, which classifies a given drug by an objective diagnostic or deciding rule. The objective of a dichotomous result requires a clear cut PM, which assigns a drug in one of the two classes "pyrogenic for humans" and "non-pyrogenic for humans". Since a threshold pyrogen value will be used, a one-sided test is appropriate for the task. Because the data are normalised by a ln-transformation, a t-test is chosen. Although the variances over the range of concentration converge by the transformation, the assumptions of equal variances do generally not hold true, because it depends on additional covariates. Therefore, the one sided Welch-t-test (Snedecor and Cochran, 1989) is applied. Due to the safety aspect of the basic problem, the hypotheses of the test are

$$H_0 : \mu_{S_i j} > \mu_{S+} \quad vs \quad H_1 : \mu_{S_i j} < \mu_{S+},$$

where $\mu_{\dots}$ denotes the parameter of location of the respective ln-transformed distribution. This approach controls the probability of false positive outcomes directly by means of its significance level α, which is chosen as 0.01, because is assumes hazard, respectively pyrogenicity, of the tested drug in H_0, and assures safety, i.e. non-pyrogenicity. The test statistic is

$$T_{S_i j} = \frac{\bar{x}_{S+} - \bar{x}_{S_i j}}{\sqrt{\frac{s^2_{S+}}{n_{S+}} + \frac{s^2_{S_i j}}{n_{S_i j}}}}.$$

The PM is built by means of the outcome of the test. Let 0 denote safety and 1 denote hazard. The classification of S_i-j is then determined by

$$S_ij = 0, \text{ if } T_{S_ij} > t_{0.99;n_{S+}+n_{S_ij}-2},$$

$$S_ij = 1, \text{ else,}$$

where $t_{0.99;n_{S+}+n_{S_ij}-2}$ the 0.99-quantile of the t-distribution with $n_{S+} + n_{S_ij} - 2$ degrees of freedom. The number of replicates for every control and sample, i.e. $n_{...}$, was harmonised to be four. Due to the possibility of removing one observation by the outlier test, the number of replicates could be reduced to three. The classification of a version of a drug is regarded as an independent decision. Therefore, the niveau α is local.

Finally, the classifications of the drugs will be summarised in 2x2 contingency table (table 3). From these tables, estimates of the sensitivity (S_E), i.e. the probability of correctly classified positive drugs and specificity (S_P), i.e. the probability of correctly classified negative drugs, will be obtained by the respective proportions. Where

$$S_E = a / (a + c) * 100\%$$
and
$$S_P = d / (b + d) * 100\%.$$

Furthermore, these estimates will be accompanied by confidence intervals, which will be calculated by the Pearson-Clopper method (Clopper & Pearson, 1934). For example, let $\hat{p}_{SE}$ denote the proportion, namely the sensitivity, under investigation. Then the confidence interval to a niveau α is calculated as

$$\left[p_{SE}^{L} = \frac{\mathrm{a}F_{2\mathrm{a};2(n_{1\cdot}-\mathrm{a}+1);\frac{\alpha}{2}}}{n_{1\cdot} - \mathrm{a} + 1 + \mathrm{a}F_{2\mathrm{a};2(n_{1\cdot}-\mathrm{a}+1);\frac{\alpha}{2}}}; p_{SE}^{U} = \frac{(\mathrm{a}+1)F_{2(\mathrm{a}+1);2(n_{1\cdot}-\mathrm{a});1-\frac{\alpha}{2}}}{n_{1\cdot} - \mathrm{a} + (\mathrm{a}+1)F_{2(\mathrm{a}+1);2(n_{1\cdot}-\mathrm{a});1-\frac{\alpha}{2}}} \right],$$

where $F_{...}$ denotes the respective quantile of the F-distribution and $n_{1\cdot}$ is the sample size of the positive drugs and a the number of correctly classified drugs.

By contaminating the drugs artificially and by defining a threshold value, which is assumed to be appropriate, the class of a drug is determined beforehand. The versions of drugs, which are effectively contaminated, but below the threshold dose, are considered to be negative, respectively safe, because their contamination is not crucial for humans in terms of ELC.

Reproducibility:

The analysis of the intra- and interlaboratory reproducibility was assessed from the three identical and independent runs conducted in each of 3 laboratories. The comparison of the three runs was carried out blindly such that the testing facility did not know the true classification of the sample, either pyrogenic or non-pyrogenic. By this procedure only

the randomness and reproducibility of the methods was assessed and not systematic errors, which may have arisen from other sources, e.g. logistics or preparation of the samples. This means that if a sample was (mis)classified in all three runs the result is reproducible regardless of the (mis)classification of the sample. Therefore, a measure of similarity, i.e. complete simple matching with equal weights, was preferred to the coefficient of correlation for 2x2 contingency tables.
The study was designed as follows: each laboratory had to conduct three independent runs with the same 12 samples (3 test items with 4 blinded spikes each) and two controls, i.e. saline as a negative control (C-) and a 0.5 EU/ml LPS-spike in saline as a positive controls (C+). The samples were derived from the three substances Gelafundine, Haemate and Jonosteril. Per run, each substance was blindly spiked twice with saline, once with 0.5 EU/ml LPS and once with 1 EU/ml LPS, which resulted in a balanced design with regard to positive and negative samples, i.e. samples expected to be pyrogenic and non-pyrogenic, respectively.

The three independent runs per testing facility provide the information on which the assessment of the intralaboratory reproducibility is based. The combined results of the three runs per testing facility were used to determine interlaboratory reproducibility. The correlation of the prediction (in terms if the Bravais-Pearson coefficient of correlation) between all runs is calculated, independent of whether that classification is true or false. A BP-correlation of 1 is calculated, if two runs gave exactly the same predictions for the twelve substances. If one run gives adverse classifications for all substances than the other, the correlation is –1. As these calculations do not need information of the true status of a sample, they were carried out blinded.

5.4 Tabulated results

Provide a summary, in graphic or tabular form, of the results.
See tables 5.4.1 and 5.4.2.

Table 5.4.1: Results of testing 3 substances 3 times by 3 laboratories. Classifications after applying the prediction model (compare to fig. 5.2.5)

Sample	DL (Konstanz)			NL 1 (Bern)			NL 2 (Oslo)		
	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3
G-0 (1)	0	0	0	0	-	0	0	0	0
G-0 (2)	0	0	0	0	-	0	0	0	0
H-0 (1)	0	0	0	0	0	0	0	0	0
H-0 (2)	0	0	0	0	0	0	0	0	0
J-0 (1)	0	0	0	0	0	-	0	0	0
J-0 (2)	0	0	0	0	0	-	0	0	0
G - 0.5	1	1	0	0	-	1	1	1	0
H - 0.5	1	1	1	1	1	1	0	0	0
J - 0.5	1	0	0	1	1	-	0	0	0
G - 1	1	1	1	1	-	1	1	1	1
H - 1	1	1	1	1	1	1	0	0	0
J - 1	1	1	1	1	1	-	1	1	1

"0"denotes "non-pyrogenic"; "1" denotes "pyrogenic".; - denotes test invalid

Table 5.4.2: Results of the validation study of 10 drugs, spiked with WHO-LPS at 0.0, 0.25, 0.5, 0.5 and 1.0 IU/ml, respectively and tested in 3 different laboratories. Samples and spikes were blinded. Classifications after applying the prediction model (compare to fig. 5.2.7).

drug (code)	**spike**		**results**		
	EU/ml	**"truth"**	Konstanz	PEI	Oslo
Beloc (BE)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**0**
	0.50	**1**	**0**	**1**	**0**
	0.50	**1**	**0**	**1**	**0**
	1.00	**1**	**0**	**1**	**1**
Binotal (BI)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**0**
	0.50	**1**	**1**	**0**	**0**
	0.50	**1**	**1**	**1**	**0**
	1.00	**1**	**1**	**1**	**NA**
Ethanol 13% (ET)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**0**
	0.50	**1**	**1**	**0**	**1**
	0.50	**1**	**1**	**0**	**1**
	1.00	**1**	**1**	**1**	**1**
Fenistil (FE)	**0.00**	**0**	**0**	**1**	**0**
	0.25	**0**	**0**	**1**	**0**
	0.50	**1**	**1**	**1**	**0**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
Glucose 5% (GL)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**0**
	0.50	**1**	**1**	**0**	**0**
	0.50	**1**	**1**	**0**	**1**
	1.00	**1**	**1**	**1**	**1**
"Drug A" 0.9% NaCl (LO)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**0**
	0.50	**1**	**0**	**0**	**0**
	0.50	**1**	**0**	**NA**	**0**
	1.00	**1**	**1**	**1**	**1**
MCP (ME)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**0**
	0.50	**1**	**1**	**1**	**0**
	0.50	**1**	**1**	**1**	**0**
	1.00	**1**	**1**	**1**	**1**
"Drug B" 0.9% NaCl (MO)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**0**
	0.50	**1**	**1**	**1**	**0**

drug (code)	spike EU/ml	"truth"	results Konstanz	PEI	Oslo
	0.50	1	1	1	1
	1.00	1	1	1	1
Orasthin (OR)	0.00	0	0	0	0
	0.25	0	0	1	0
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Sostril (SO)	0.00	0	0	NA	0
	0.25	0	0	1	0
	0.50	1	0	1	1
	0.50	1	0	1	1
	1.00	1	1	1	1

"0"denotes "non-pyrogenic"; "1" denotes "pyrogenic"; NA is not assessed.

5.5 Coding of data

For each set of data, indicate whether coded chemicals were tested, whether experiments were conducted without knowledge of the chemicals being tested, and the extent to which experiments followed GLP guidelines.

Blinding of drugs and/or spikes is indicated with the data.

5.6 Circumstances

Indicate the "lot-to-lot" consistency of the test substances, the time frame of the various studies, and the laboratory in which the study or studies were conducted. A coded designation for each laboratory is acceptable.

In each part of the study, all samples are derived from one (clinical) lot.

5.7 Other data available

Indicate the availability of any data not submitted for external audit, if requested.

All relevant data were submitted with the present BRD.

6 Test Method Accuracy

6.1 Accuracy

Describe the accuracy (e.g., concordance, sensitivity, specificity, positive and negative predictivity, false positive and negative rates) of the proposed test method compared with the reference test method. Explain how discordant results in the same or multiple laboratories from the proposed test were considered when calculating accuracy.
Test method accuracy was assessed in two large scale experiments performed with the drugs outlined in table 3.3.1 and table 3.3.2 in section 3 respectively. As described before one experiment was performed in an early stage of the study with 3 different drugs, tested 3 times and the other final experiment all drugs were tested once in the three participating laboratories. From the first experiment a preliminary estimate of sensitivity and specificity can be figure out, whereas the second is regarded as the established accuracy for the WB/IL-1 assay.

6.1.1 Preliminary estimate of the accuracy of the WB/IL-1 test. In an early stage of the study a different concept for interference testing was used. The developing laboratories determined for each drug (outlined in table 3.3.2, section 3.3) the smallest dilution within the MVD that showed no interference or an acceptable degree of interference with the spike recovery. In general the lowest dilution of the sample allowing for a 50-200% spike recovery was chosen. In addition, the positive control (PC) set at 0.5 EU/ml saline was used as the classification threshold. The laboratory procedure as described in the method protocol was maintained throughout the study. Although it was realized there were some drawbacks to the concept for interference testing and applying the PC as a threshold, this small scale study allows for a preliminary estimate of the accuracy of the WB/IL-1 method.
It has to be noted that this part of the study was designed to provide an estimate of the intra- and inter laboratory reproducibility. Therefore it will also be discussed in detail in section 7 (Test Method Reliability).

According to the PM applied during an early phase of the study the outcome (positive/negative) is related to the positive control (PC=0.5 EU/ml). If absorbance of sample > absorbance of PC, then the sample is classified as positive. If absorbance of sample < PC, then the sample is classified as negative. While performing the experiments during this phase it was realized that there is a flaw between interference testing and the PM. After the interference testing, the lowest dilution of the sample allowing for a 50-200% spike recovery was chosen. It is obvious that even with a flawless repeatability of the assay; a spike recovery between 50%-100% would be classified as negative according to the preliminary PM. In addition, due to unforeseen problems with the preparations of the spike, the recovery of the spikes was far below 100%. (This is outside the scope of the study and will not be discussed). As a consequence of the employed preliminary setup of the study the sensitivity will be underestimated, and the specificity will be overestimated.

In short, three test substances were spiked with WHO-LPS at four levels (0, 0, 0.5 and 1.0 EU/ml respectively), which resulted in 12 samples with a balanced design with regard to positive and negative samples (i.e. samples expected to be pyrogenic and non-pyrogenic respectively. These 12 sample were three times tested in three laboratories. In total there were 108 classifications from 12 samples in 3 runs and in 3 laboratories (3x3x12=108). Results are described in detail in section 7. A 2x2 contingency table was constructed (table 6.1.1), from which the estimates of sensitivity and specificity can easily be derived.

Table 6.1.1: 2x2 contingency table. The prediction model applied to a preliminary study.

		True status of samples		**Total**
		+	-	
PM	+	36	0	36
	-	14	50	64
	Total	50	50	100

The specifications of specificity and sensitivity described in section 5.3 were applied to these results and the specificity (Sp) of the WB/IL-1 assay is 100% (50/(50+0)*100%), 95% confidence interval [0.929;1]. The sensitivity (Se) equals 72% (36/(36+14) *100%), 95% confidence interval [0.575; 0.838]. As outlined previously the specificity is overestimated and the sensitivity is underestimated as a result of the design of this part of the study.

6.1.2 Test method accuracy of the proposed WB/IL-1 method. To assess accuracy of the proposed method, 10 substances (listed in table 3.1.1, section 3) were spiked with five different concentrations of the WHO-LPS (one of which is negative). Thus, in total, 50 samples have been tested in each laboratory.
To permit the application of the chosen PM, each drug was diluted to its individual MVD, which has been calculated using the ELC to that drug (listed in section 3). Lesser dilutions were tested by the DL, and showed no interference. Therefore interference was not expected at the individual MVD. Each substance had their own specific set of endotoxin spike solution, ensuring that after spiking the undiluted drug, it contained 0.0; 0.25; 0.5; 0.5 and 1.0 EU/ml respectively when tested at its MVD. A detailed description of the sample preparation was supplied to the three independent laboratories (shown in table 5.1.1 for convenience). To put more weight to the result of this part of the validation, the spikes were blinded and coded by QA ECVAM. The raw data and the graphical presentation of these raw data are shown in the section 5 (table 5.4.2). Accuracy was assessed by applying the PM to the results (summarized in table 5.3.2) and evaluating the concordance in this section in a two by two contingency table (table 6.1.2). As described above 10 substances, spiked with 5 different WHO-LPS concentrations were tested in three laboratories and consequently a maximum of 150 data were available for analysis.

As intralaboratory reproducibility was successfully shown in previous experiments (analyzed in section 7), only one run performed in each laboratory was considered sufficient.

Table 6.1.2: 2x2 contingency table. Prediction model applied to the WB/IL-1 test result of 10 different substances assessed in three different laboratories.

		True status of samples +	-	Total
PM	+	64	4	68
	-	24	55	79
	Total	88	59	147

Of the 150 available data, only three sets of 4 replicates did not comply with the quality criteria as defined in the method protocol (CV <0.45) and were removed from the analysis. The specificity and sensitivity of the WB/IL-1 method could be estimated as described in section 5.3.

The specificity of the WB/IL-1 assay is 93.2% (55/(55+4)*100%), 95% confidence interval [0.883;0.996]. The sensitivity equals 72.7% (64/(64+24) *100%), 95% confidence interval [0.622;0.817]. (See table 6.1.3). The specificity varied from 78.9% up to 100% within the three laboratories, and the sensitivity varied from 62.1% up to 100%.

Table 6.1.3: Specificity and sensitivity of the WB/IL-1 assay

	N total	N correctly identified	proportion	95% CI lower limit	95% CI upper limit
Specificity (Sp)	59	55	93.2%	88.3%	99.6%
Sensitivity (Se)	88	64	72.7%	62.2%	81.7%

6.2 Concordancy to *in vivo* reference method

Discuss results that are discordant with results from the in vivo reference method.
Not applicable.

6.3 Comparison with reference methods

Discuss the accuracy of the proposed test method compared to data or recognized toxicity from the species of interest (e.g., humans for human health-related toxicity testing), where such data or toxicity classification are available. This is essential when the method is measuring or predicting an endpoint for which there is no preexisting method. In instances where the proposed test method was discordant from the in vivo reference test method, describe the frequency of correct predictions of each test method compared to recognized toxicity information from the species of interest.
Not applicable.

6.4 Strength and limitations

State the strengths and limitations of the proposed test method, including those applicable to specific chemical classes or physical-chemical properties.
It appears the proposed test is applicable to most classes of medicinal products, at least those that are non- or low-toxic to cells *in vitro*. In addition, the test may be employed to

assess pyrogenicity of various medical devices, such as (biological) bovine collagen bone implants.

6.5 Data interpretation

Describe the salient issues of data interpretation, including why specific parameters were selected for inclusion.
No issues.

6.6 Comparison to other methods

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the results obtained with both test methods should be compared with each other and with the in vivo reference test method and/or toxicity information from the species of interest.
Not applicable.

7 Test Method Reliability (Repeatability/Reproducibility)

7.1 Selection of substances

Discuss the selection rationale for the substances used to evaluate the reliability (intralaboratory repeatability and intra- and interlaboratory reproducibility) of the proposed test method as well as the extent to which the chosen set of substances represents the range of possible test outcomes.

The rationale for the selection of the substances is described in section 3.3. In short: for the present studies endotoxin (WHO-LPS) was selected as the model pyrogen, since it is well defined biological standard and readily available. Selected test substances were medicinal products available on the market. These batches are released by the manufacturers and comply with the Marketing Authorisation file and European Pharmacopoeia. Therefore these batches are considered to contain no detectable pyrogens. To test the method reliability the medical products were spiked with endotoxin.

7.2 Results

Provide analyses and conclusions reached regarding the repeatability and reproducibility of the proposed test method. Acceptable methods of analyses might include those described in ASTM E691-92 (13) or by coefficient of variation analysis.

In an early phase of the study, the intralaboratory repeatability and reproducibility of the test method was assessed in a series of experiments conducted in the DL. Series of blanks (saline, 0 EU/ml) and spikes of WHO-LPS in saline were tested. These experiments (1A, 1B, 2A, 2B and 2C) are summarized in table 7.2.1.

Table 7.2.1: Summary of experiments with WHO-LPS in saline.

Experiment	Spikes (EU/ml) in saline	n (per spike)	Repetitions of experiment	N
1A	0; 0.5	32	1	64
1B	0; 0.063; 0.125; 0.25; 0;5	12	1	60
2A	0; 0.5	12	3	72
2B	0; 0.25; 0.5	8	3	72
2C	0; 0.5	5	8	80

The data were used to answer questions regarding the nature of the distribution, the variance and its behavior over the range of response in replicated measurements under identical conditions. In addition reliability of the test method was assessed by the maintenance of minimum criteria, e.g. a detection limit below an a-priori chosen positive control or a dose-dependent standard curve (table 7.2.1, experiment 1B). With the data of this experiment an assessment of the limit of detection of the test can be done by calculating the smallest spike, which can be discriminated from the blank.

The second group of experiments was meant to analyze the variation in detail. For this purpose the major sources of variation were assessed separately, i.e. donor (experiment 2A) and operator (exp. 2B). A total of 348 data were collected and analyzed.

First the shape of the distribution at a spike was assessed (not shown). Most of the data showed normal-distribution.
Based on the experience that there is a monotone increasing relationship between the mean-responses and the variation (empirical variance or standard deviation), the analysis focuses on the coefficient of variation (CV). The CV should be distributed symmetric around a constant factor, if the mean-variance relationship is linear. A plot of all CVs against their corresponding means is shown in figure 5.2.1. From the figure it is clear that at this stage of the study, the CV for the blank is in six of 16 cases high. For the other spikes, first of all the 0.25 and the 0.5 EU/ml-spike, the variation for the sets of replicates is low. As only WHO-LPS was examined up to this point, it was envisaged that the CV would increase with other substances being tested. For CV criteria applied as a validity criteria of the WB/IL-1 assays was arbitrarily set at CV<0.25.

The outliers were identified on the assumption of normally distributed data and this allows to apply a parametric test. At this point the Grubbs-test was chosen and the kind of outlier (lower or upper) and the significance level α were recorded. Altogether there were 2 lower outliers and 10 upper outliers (equally divided between 5% and 1% significance level). Overall the amount of outliers is about 3%. The outliers were located all over the ELISA-plates and there was no obvious scheme. In addition, the raw data (plate-readouts) showed no obvious edge effects or trends.

The results of test 1A and 1B (figure 5.2.2) show that (after removal of outliers in the blank) the 0.25 EU/ml of spike can be discriminated statistically from the blank and the highest spike (0.5 EU/ml) can be detected easily.

Test 2A was designed to assess the behavior of a donor in time. The blood was taken on three successive days. Data are presented in figure 5.2.3. The donor showed in general the same behavior on the three days. The small deviations in the height of the response are negligible. Hence, the variation of a suitable donor is low and is considered to be no critical issue in the WB/IL-1 assay.

Three operators in parallel, using blood from the same donor conducted experiment 2B (figure 5.2.4). In general, the data are similar, but it is obvious that the sensitivity of the assay seems to depend on the operator. But still the data of the 0.25 EU/ml spikes can easily be discriminated from their corresponding blanks.

The final experiment was designed to show the robustness of the assay with respect to different donors. Therefore 8 donors were involved and for each donor five replicates of each of the spikes (0; 0.5EU/ml) were generated. Data are presented in figure 5.2.5. For donors 5-8 the variability within the blank-replicates were high. Some variation in sensitivity for LPS between the donors is obvious, especially donor 5 shows a lower response. But every donor reacts to the 0.5 EU/ml-spike. This experiment reveals that there is a certain effect of the covariate "donor" which is however not crucial to a qualitative PM.

In conclusion: The most critical issue identified is the variation within the sets of blanks, but this is probably caused by the handling of the assay. The WB/IL-1 assay is robust against all examined variables. Although the experiments revealed an effect for the covariates "blood donor", "operator" and "day", the sensitivity of the assay is at least 0.25 EU/ml for all experiments, thus 0.5 EU/ml is always detectable. Therefore the intralaboratory repeatability is considered satisfactory. The 3% percentage outliers, as determined by the Grubbs test is considered acceptable. The validity criteria of the WB/IL-1-assay as recorded in the method protocol, are based on these experiments, i.e. CV< 0.25, lower limit of detection 0.25 EU/ml.

Intra- and inter laboratory reproducibility.
After transfer of the WB/IL-1 assays to two other laboratories, a dose response experiments was performed by all three laboratories. For this study 6 or 7 concentrations were tested in a dose response curve (typically 0, 0125, 0.25, 0.5, 1, 2, 4 EU/ml, at least 8 replicates). A participating laboratory qualified for taking part in next part of the study by producing a dose response curve, with a limit of detection of at least 0.25 EU/ml and a CV < 0.25 (data not shown.).

The intra- and interlaboratory reproducibility was assessed by testing 3 different medicinal substances, Gelafundin, Jonosteril and Haemate (described in table 3.3.2, section 3.3.). Test substances and their spikes were appropriately blinded. Test substances were tested, at a predefined dilution above the MVD, independently in 3 different laboratories, 3 times each. The three test substances were spiked with WHO-LPS at four levels (0, 0, 0.5 and 1.0 EU/ml respectively), which resulted in 12 samples with a balanced design with regard to positive and negative samples (i.e. samples expected to be pyrogenic and non-pyrogenic respectively. In addition a negative control (saline) and positive control (0.5 EU/ml in saline) were included to establish assay validity. To avoid interference, the DL performed interference testing in terms of the BET, i.e. 50-200% spike recovery, and decided on the dilution of the test substances. Dilutions chosen for Gelafundine, Haemate, Jonosteril were 1:2, 1:20 and 1:2 respectively. The data derived by the RIVM are taken as an example of the three laboratories. The raw data and a graphical presentation of the absorbance values are shown in section 5 (raw data exp.5 and fig. 5.2.5).

From the experiment with LPS-WHO only it was concluded that CV for the WB/IL-1 assay is < 0.25, which is acceptable. It was envisaged that the CV was likely to be higher when testing different substances (different matrices) and was assessed for the current set of data. A plot of all CVs for all sets of 4 replicates of a drug with a spike is shown in figure 5.2.7. From the figure it is clear that the CV for a set of 4 replicates of one spike concentration is usually below 0.45, which is considered acceptable for a biological assay. Only one set of data showed an exceptional high (CV>1.1) which is probably due to a pipetting error. For the remainder of the studies the CV criteria applied as validity criteria of the WB/IL-1 assays was arbitrarily set at CV<0.45.

The analysis of the intralaboratory reproducibility was assessed from the three identical and independent runs conducted in each laboratory. The comparison of the three runs was

carried out blindly such that the laboratory did not know the true classification of the sample (either pyrogenic or non-pyrogenic). By this procedure only the randomness and reproducibility of the methods was assessed and not systematic errors, which may have arisen from other sources, e.g. logistics or preparation of the samples. This means that if a sample was misclassified in all three runs the result is 100% intralab reproducible (regardless of the misclassification of the sample).

According to the preliminary PM applied during this phase of the study the outcome (positive/negative) was related to the positive control (PC=0.5 EU/ml). If absorbance of sample > absorbance of PC, then sample is classified as being positive. If absorbance of sample < PC, sample is classified as negative (positive/pyrogenic = 1, negative/non-pyrogenic = 0).
During this phase it was realized that there is a flaw between interference testing and the PM. After the interference testing, the lowest dilution of the sample allowing 50-200% spike recovery was chosen. It is obvious that even with a flawless repeatability of the assay, a spike recovery between 50%-100% would be classified as negative according to the preliminary PM.

From the three indepent runs summarized in table 5.4.1, the intralaboratory reproducibility can be calculated for the separate laboratories (table 7.2.2). For these calculations there is no need for information of the true status of the sample. A minimum criterion for the establishment of an assay is that experiment carried out with the same samples should result in a high concordance of classifications.
For NL1 the calculations were limited, because the sensitivity criterion, i.e. a significant difference between C- and C+, was not fulfilled for Gelafundine in run 2 and Jonosteril in run 3. This results in only 28 samples for NL1 instead of 36.
From table 7.2.2 it can be read that the intralaboratory reproducibility is very good (89 – 94%) for all three participating laboratories.

Table 7.2.2 : Intralaboratory reproducibility, assessed by correlation between different runs. Result of testing 3 substances 3 times by 3 laboratories.

	DL (Konstanz)	**NL1 (Berne)**	**NL2 (Oslo)**
Run 1 - Run 2	92% (11/12)	100% (8/8)	100% (12/12)
Run 1 - Run 3	83% (10/12)	88% (7/8)	92% (11/12)
Run 2 - Run 3	92% (11/12)	NA	92% (11/12)
Mean	89%	94%	94%
Proportion showing the same result in 3 runs	83%	NA	92%

NA = not assessed for lack of sufficient data

BRD: WB/IL-1 March, 2006

The interlaboratory reproducibility of the WB/IL-1 method was assessed in a similar manner to the intralaboratory reproducibility. A summarizing method to combine the three runs per laboratory is considered not appropriate, because it would mask misclassification. Therefore each run of one laboratory was compared with all runs of another laboratory. This results optimally in 108 comparisons between the data sets of two laboratories. The measure of similarity is then the proportion of equally classified samples. These proportions are summarized in table 7.2.3, show that there is a good interlaboratory reproducibility of at least 68%.

Table 7.2.3: Interlaboratory reproducibility. assessed by interlaboratory correlations. Result of testing 3 substances 3 times by 3 laboratories.

Laboratories	Interlaboratory reproducibility	Number of equal predictions
DL – NL1	92%	77 / 84
DL – NL2	77%	83 / 108
NL1 – NL 2	68%	57 / 84
Mean	79%	

DL = Konstanz; NL1 = Bern; NL2 = Oslo

Also from the result of the large scale study (testing 10 substances spiked with 5 separate spikes), the interlaboratory reproducibility can be estimated (table 7.2.4). All the samples were correctly identified by one of the laboratories (DL). The reproducibility varied from 70% to 82% between two laboratories. All three laboratories found the same result for 27 out of 47 samples (57%).

Table 7.2.4: Interlaboratory reproducibility: Assessed by testing of 10 substances, spiked 5 times. One run of 50 samples by three different laboratories.

Laboratories	Interlaboratory Reproducibility	Number of equal predictions
DL - NL1	73%	35 / 48
DL - NL2	82%	40 / 49
NL1 – NL2	70%	33 / 47
Mean	75%	
same result in all laboratories	57%	27 / 47

DL = Konstanz; NL1 = BFL, Germany NL2 = Oslo

Conclusion: It is shown that the intralaboratory reproducibility, assessed by the proportion of equally classified samples between different runs varies from 89% to 94% between the three participating laboratories. The interlaboratory reproducibility between two laboratories varied from 68% to 92% in one large scale blinded experiment and from 70% to 82% in the other large scale blinded experiment. All three participating laboratories predicted the same in respectively 79% and 57% of the measurements. It has

to be noted that a substantial part of the samples was 0.5 EU/ml, at or close to the arbitrary demarcation point of the WB/IL-1 assay

7.3 Historical data

Summarize historical positive and negative control data, including number of experiments, measures of central tendency, and variability.
Not applicable.

7.4 Comparison to other methods

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the reliability of the two test methods should be compared and any differences discussed.
Not applicable.

BRD: WB/IL-1 March, 2006

8 Test Method Data Quality

8.1 Conformity

State the extent of adherence to national and international GLP guidelines (7-12) for all submitted data, including that for the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method. Information regarding the use of coded chemicals and coded testing should be included.
The studies were done in accordance to the guidelines for GLP. Written protocols and approved standard operating procedures were followed during the entire course of the study. Deviations were recorded and, where appropriate, approved in amendments. All data are stored and archived. As mentioned, samples were appropriately blinded.

8.2 Audits

Summarize the results of any data quality audits, if conducted.
No audits were conducted.

8.3 Deviations

Discuss the impact of deviations from GLP guidelines or any noncompliance detected in the data quality audits.
Not applicable.

8.4 Raw data

Address the availability of laboratory notebooks or other records for an independent audit. Unpublished data should be supported by laboratory notebooks.
All records are stored and archived by the contributing laboratories and available for inspection.

9 Other Scientific Reports and Reviews

9.1 Summary

Summarize all available and relevant data from other published or unpublished studies conducted using the proposed test method.

Relevant data obtained with the proposed method are described in a number of published studies and reports. The most important ones for this BRD are included in the Appendix B as hardcopies and referenced in Section 12, whereas for others only the references are given in section 14. In most of the study reports the WB/IL-1 is named *in vitro* pyrogen test or IPT.

The establishment of the whole blood test as an alternative to the rabbit pyrogen test as well as the comparison to the BET is described below.

Further applications were developed by adaptation to the basic whole blood test e.g. to measuring pyrogenic contaminations of medical devices and measuring the air quality in the working place and references are included in Section 14 in part 2 and 3.

A total of 96 batches of parenteral pharmaceuticals from 21 indication groups were tested using the WB/IL-1 test and compared to data from the rabbit and BET, if available (Jahnke et al, 2000). For these batches of parenteral drugs it was shown that the result of the three methods correlated well. In one case (an amino acid-containing infusion solution) a pyrogen-containing batch was clearly detected by all three testing systems. The other parenteral pharmaceuticals remained negative in all assays. It is worth mentioning that all the products could be tested in the whole blood test, in some cases after interfering factors had been excluded. A few drugs (e.g. dopamine) were found to affect the sensitivity of the WB/IL-1 assay and hence caused interference, but this could be overcome by dilution of the drug.

In a preliminary study (Fennrich et al., 1999), the suitability of the test was tested by determining the LPS retrieval in spiked pharmaceutical samples at the border line concentrations given in the Ph. Eur. for endotoxins (ELC), which should be detectable also using the human WB/IL-1 test.

Human serum albumin belongs to those substances that still are tested in the rabbit pyrogen test. Spreitzer et al (2002) compared the sensitivity of the rabbit assay with the WB/IL-1 assay using 29 defined human albumin samples: plain, spiked with 5 EU/ml and 10 EU/ml respectively. The unspiked samples were negative in both assays. Both the borderline 5 EU/kg and the 10 EU/kg partially led to results of the rabbit test (conducted with 3 rabbits), which would cause further testing with additional animals. In contrast, the human whole blood assay resulted in a 100% detection for the 5 EU/ml and 10 EU/ml endotoxine spike. The human whole blood test resulted in at least the same level of security for the products as the rabbit pyrogen test did. After further dilution of the 29 spiked albumin samples to contain 0.5 EU/ ml, 18 samples were still positive in the WB/IL-1 assay but there were 11 negative results too.

In a manuscript of Schindler et al. (2003) the reactivity of human and rabbit blood *in vitro* towards Gram negative and Gram-positive stimuli were compared directly using an

in vitro whole blood test (endpoint; IL-1) for both species. The reactivity of the two species towards LPS was found to be similar, whereas human blood was more sensitive for LTA (lipoteichoic acid) than rabbit blood. The results suggested that the test with human blood to detect contaminations in e.g. parenteral drugs, might predict the human reaction to real life contamination better than the rabbit pyrogen test.

A Gram-positive standard derived form B. subtilis has been developed by the same research group (Konstanz University) and was reviewed in numerous different articles. This lipoteichoic acid, is BET negative which however reacts positive in the WB/IL-1 assay. Identification, isolation and purification of other Gram-positive stimuli are subject of ongoing research.

It is stressed throughout these studies using whole blood that only healthy donors not taking any medication must be used for testing pharmaceuticals. Various drugs such as cortisone suppress interleukin release and would therefore exclude a blood donor from the experiment. An infection can stimulate release, as reflected in an increased baseline value or an abnormally high interleukin response. Therefore, the WB/IL-1 test may only be used if samples have first been shown not to cause interference. The blood group of the human donors does not influence the results of the assay.

9.2 Discussion

Comment on and compare the conclusions published in independent peer-reviewed reports or other independent scientific reviews of the proposed test method. The conclusions of such scientific reports and reviews should be compared to the conclusions reached in this submission. Any ongoing evaluations of the proposed test method should be described.

The validation study summarised in this BRD is the first, which extensively addresses specificity and accuracy using actual medicinal products spiked with endotoxin. Hence, there are no comparing reports in independent peer-reviewed journals available. However, the validation study confirms conclusions of scientific reports. E.g. several preliminary studies (e.g. Jahnke et al.2000, Fennrich et al. 1999, Spreitzer et al 2002) showed that the WB/IL-1 assay is suitable to test different types of pharmaceuticals. This finding is confirmed by the current validation study, where 11 different pharmaceuticals were tested. In addition, both studies indicate that (pyrogen free) batches which passed the current batch release scheme and are available on the market, show rarely a false positive reactivity in the WB/IL-1 assay. Jahnke's study was conducted by an experienced laboratory, whereas relatively inexperienced laboratories were also involved in the validation study. This may account for the less than 100% specificity in the validation study.

Finally, Charles River Endosafe offers the whole blood test under the name IPT (*In vitro* Pyrogen Test) worldwide in a highly standardized kit-version. Frequent symposia and workshops with coworkers of Charles River together with the University of Konstanz take place in order to train interested parties and introduce the IPT to users. The introduction and optimization of cryopreserved human whole blood is expected to overcome all final obstacles to standardization.

9.3 Results of similar validated method

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the results of studies conducted with the validated test method subsequent to the ICCVAM evaluation should be included and any impact on the reliability and accuracy of the proposed test method should be discussed.

As mentioned, *in vitro* monocyte activation test methods for the detection of pyrogenic contaminants are being developed over the course of the past two decades. A number of variants have been described, although the underlying principle of each variant remains the same. The test preparation is cultured with monocytoid cells, either as peripheral blood mononuclear cells, PBMC, (diluted) whole blood or cells of a monocytoid cell line such as MM6. Accuracy and specificity of these test methods are comparable, but in general methods using whole blood, PBMC and the MM6 cell line appear to perform best (Hoffmann et al, 2005b; summarized in table 9.3.1).

Table 9.3.1 summarises the performance of *in vitro* methods presented in the five BRDs and Table 9.3.2 compares the *in vivo* and *in vitro* pyrogen tests regarding their strengths, weaknesses, costs, time, limitations.

However, most studies (as this one) are done with model pyrogens and as yet little experience is available in the field, e.g. as part of the final batch release test-package. Experience and thus confidence in these methods will grow once regulatory authorities approve these methods and more manufacturers start to employ them. Then, on a case by case situation, it should be determined which method is best suited for the actual situation and demonstrates to pick out the appropriate, i.e. pyrogenic batches of the medicinal product.

BRD: WB/IL-1 March, 2006

Table 9.3.1: Summary of the performance of in vitro pyrogen tests based on monocytoid cells (see Tables 7.2.2; 7.2.4; 6.1.3)

Test	System	Read-out	Intralaboratory reproducibility (%)	Interlaboratory reproducibility (%)	Sensitivity (%)	Specificity (%)
WB/IL-6	whole blood	IL-6	DL: 83.3 NL1: 94.4 NL2: 100	DL-NL1: 85.4 DL-NL2: 85.4 NL1-NL2: 92.0	88.9	96.6
WB/IL-1	whole blood	IL-1β	DL: 88. 9 NL1: 95.8 NL2: 94.4	DL-NL1: 72.9 DL-NL2: 81.6 NL1-NL2: 70.2	72.7	93.2
96-wells WB/IL-1 [1]	whole blood	IL-1β	-	DL-NL1: 88.1 DL-NL2: 89.7 NL1-NL2: 91.5	98.8	83.6
CRYO WB/Il-1	cryo whole blood	IL-1β	-	DL-NL1: 91.7 DL-NL2: 91.7 NL1-NL2: 91.7	97.4	81.4
KN CRYO WB/Il-1 [2]	cryo whole blood	IL-1β	-	DL-NL1: 83.3 DL-NL2: 100 NL1-NL2: 83.3	88.9	94.4
PBMC/IL6	PBMC	IL-6	DL: 94.4 NL1: 100 NL2: 94.4	DL-NL1: 84.0 DL-NL2: 86.0 NL1-NL2: 90.0	92.2	95.0
PBMC-CRYO/IL-6 [3]	PBMC	IL-6	-	DL-NL1: 96 DL-NL2: 76 NL1-NL2: 80	93.3	76.7
MM6/IL-6	MM6	IL-6	DL: 100 NL1: 94.4 NL2: 94.4	DL-NL1: 90.0 DL-NL2: 89.6 NL1-NL2: 83.3	95.5	89.8

DL = developing laboratory; NL1, NL2 = naive laboratory 1 and 2
1 = data provided in Section 13 of WB/IL-1 BRD
2 = data provided in Section 13 of CRYO WB/IL-1 BRD
3 = data provided in Section 13 of PBMC/IL-6 BRD

Table amended from Hoffmann et al 2005b; results with THP cells not included

BRD: WB/IL-1 March, 2006

Table 9.3.2: Comparison of the in vivo and in vitro pyrogen tests regarding their strengths, weaknesses, costs, time, limitations

	Rabbit pyrogen test	**BET / LAL**	**In vitro pyrogen test**
Test materials	Liquids	Clear liquids	Liquids, potentially cell preparations, solid materials
Pyrogens covered	All (possible species differences to humans for non-endotoxin pyrogens)	Endotoxin from Gram-negative bacteria	(probably) all
Limit of detection (LPS)	0,5 EU	0,1 EU (some variants down to 0,01 EU)	0,5 EU (validated PM), some variants down to 0,001 EU
Ethical concerns	Animal experiment	About 10% lethality to bled horseshoe crabs	Some assays: blood donation
Costs*	High (200-600$/sample)	Low (50-150$/sample)	Medium (100-350$/sample)
Time required	27 h	45 min	24-30h**
Materials not testable	Short-lived radiochemicals, anesthetics, sedatives, analgetics, chemotherapeutics, immunomodulators, cytokines, corticosteroids	Most biologicals, glucan-containing preparations (herbal medicinal products, cellulose-filtered products), lipids, microsomes, cellular therapeutics	Not known (some of the materials not testable in rabbits require adaptations)
Others	No positive or negative control included, strain differences, stress affects body temperature	Potency of LPS from different bacterial species in mammals not reflected, false-positive for glucans	Possible donor differences, need to exclude hepatitis/HIV and acute infections / allergies of donors, dedifferentiation of cell lines

* = We consulted the laboratories participating in the validation study and a consultant regarding the costs of the tests. The figures we received vary significantly depending on the facility (e.g. industry, contract laboratory, control authority), frequency of testing, specific test requirements, country, etc.

** = interference testing might increase duration by 24 hours

10 Animal Welfare Considerations (Refinement, Reduction, and Replacement)

10.1 Diminish animal use

Describe how the proposed test method will refine (reduce or eliminate pain or distress), reduce, or replace animal use compared to the reference test method.

Depending on the medicinal product, one of two animal-based pyrogen tests is currently prescribed by the respective Pharmacopoeias, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET). The rabbit pyrogen test detects various pyrogens but alone the fact that large numbers of animals are required to identify a few batches of pyrogen-containing samples argues against its use. In the past two decades, the declared intention to refine, reduce and replace animal testing, has lowered rabbit pyrogen testing by 80% by allowing to use the BET as an alternative pyrogen test for certain medicinal products. Bacterial endotoxin is the pyrogen of major concern to the pharmaceutical industry due to its ubiquitous sources, its stability and its high pyrogenicity. With the BET, endotoxin is detected by its capacity to coagulate the amoebocyte lysate from the haemolymph of the American horseshoe crab, *Limulus polyphemus*, a principle recognised some 40 years ago (Levin and Bang, 1964). In the US, *Limulus* crabs are generally released into nature after drawing about 20% of their blood and therefore most of these animals survive. However, the procedure still causes mortality of about 30,000 horseshoe crabs per year, which adds to more efficient threats of the horseshoe crab population like its use as bait for fisheries, habitat loss and pollution.

The proposed test method is an alternative for the rabbit test and the BET. By replacing the rabbit test or the BET, the lives of rabbits and horseshoe crabs are spared.

10.2 Continuation of animal use

If the proposed test method requires the use of animals, the following items should be addressed:

10.2.1 *Describe the rationale for the need to use animals and describe why the information provided by the proposed test method requires the use of animals (i.e., cannot be obtained using non-animal methods).*

Not applicable.

10.2.2 *Include a description of the sources used to determine the availability of alternative test methods that might further refine, reduce, or replace animal use for this testing. This should, at a minimum, include the databases searched, the search strategy used, the search date(s), a discussion of the results of the search, and the rationale for not incorporating available alternative methods.*

Not applicable.

10.2.3 *Describe the basis for determining that the number of animals used is appropriate.*

Not applicable.

BRD: WB/IL-1 March, 2006

10.2.4 *If the proposed test method involves potential animal pain and distress, discuss the methods and approaches that have been incorporated to minimize and, whenever possible, eliminate the occurrence of such pain and distress.*
Not applicable.

11 Practical Considerations

11.1 Transferability

Discuss the following aspects of proposed test method transferability. Include an explanation of how this compares to the transferability of the in vivo reference test method and, if applicable, to a comparable validated test method with established performance standards.

In general, the proposed test method is not unlike other bioassays and immunoassays that are performed routinely in many laboratories.

11.1.1 *Discuss the facilities and major fixed equipment needed to conduct a study using the proposed test method*.

No extraordinary facilities are required. General laboratory equipment for aseptic operations and analytical instruments for performing immunoassays, e.g. microtiter plate reader and –washer, are sufficient to perform the proposed test method.

11.1.2 *Discuss the general availability of other necessary equipment and supplies*.

All supplies and reagents are readily available on the market. In contrast, availability of sufficient rabbits of adequate weight and in good health for the *in vivo* reference test is sometimes reported a limitation.

11.2 Training

Discuss the following aspects of proposed test method training. Include an explanation of how this compares to the level of training required to conduct the in vivo reference test method and, if applicable, a comparable validated test method with established performance standards.

11.2.1 *Discuss the required level of training and expertise needed for personnel to conduct the proposed test method.*

The proposed test method requires personnel trained for general laboratory activities in cell biology and immunochemistry or biochemistry. Techniques they should master are not unlike cell culture (aseptic operations) and immunological techniques (especially ELISA). Such expertise is available in most if not all QC-laboratories.

11.2.2 *Indicate any training requirements needed for personnel to demonstrate proficiency and describe any laboratory proficiency criteria that should be met.*

Personnel should demonstrate that they master the execution of the test. The candidate should demonstrate to meet all the appropriate assay acceptance criteria and yield accurate results (outcome) using selected test items.

11.3 Cost Considerations

Discuss the cost involved in conducting a study with the proposed test method. Discuss how this compares to the cost of the in vivo reference test method and, if applicable, with that of a comparable validated test method with established performance standards.
Three factors contribute to the cost of the proposed test method: availability of monocytoid cells, cost of the reagents for the immunoassay and, last but not least, personnel.
Since the proposed test method is relatively more labor-intensive, it is estimated that the cost of the proposed test method is more then the BET or the *in vivo* reference test using rabbits. Obviously, a higher throughput of tests (runs/year) such as in a QC-laboratory of a multi-product facility or in a Contract Research Organization will significantly reduce the costs per assay.
However, especially with pharmaceuticals of biological origin, the proposed test method may be cost-effective, since these products all to often are incompatible with the BET and by their nature preclude the reuse of the rabbits.

11.4 Time Considerations

Indicate the amount of time needed to conduct a study using the proposed test method and discuss how this compares with the in vivo reference test method and, if applicable, with that of a comparable validated test method with established performance standards.
Esssentially the test stretches two working days. On day one the testing materials are prepared and incubated overnight with the monocytoid cells. On the second day the amount of excreted cytokines is determined by immunoassay. The total time from start to result is approximately 24 hours.
It is thus concluded that the proposed test method will take more time when compared to the alternative tests, either the rabbit test or the BET. It should be noted that rabbits are tested prior to their first use by a sham test.

12 References

List all publications referenced in the submission.

References in bold are included as hardcopies in Appendix B

Barnett, V., Lewis, T. (1984). Outliers in statistical data. 2nd edition, Chichester, US: John Wiley & Sons, pp 171-172.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Clopper, C. J., Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26: 404-413

Co Tui & Schrift, M.H. (1942). A tentative test for pyrogen in infusion fluids. Proc. Soc. Exp. Biol. Med. 49: 320-330.

Cooper JF, Levin J, Wagner Jr. HN (1971). Quantitative comparison of *in vitro* and *in vivo* methods for the detection of endotoxin. J Lab Clin Med 78: 138-145.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Fennrich S, Fischer M, Hartung T, Lexa P, Montag-Lessing T, Sonntag H-G, Weigandt M und Wendel A. (1999). Detection of endotoxins and other pyrogens using human whole blood. Dev. Biol. Standards 101:131-139

Fennrich S, Wendel A and Hartung T. (1999). New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 16:146-149

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005a). Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Pertbauer A, Schindler S, et al. (2005b) Validation of novel pyrogen tests based on human monocytoid cell lines. J Immunol Methods, 298(1-2):161-173.

Hothorn, L. A. (1995). Biostatistical analysis of the 'control vs. k treatments' design including a positive control group. In: Testing principls in clinical and preclinical trials, pp 19-26, Stuttgart, Germany: Gustav Fischer Verlag.

Jahnke M, Weigand, Sonntag H-G. (2000). Comparative testing for pyrogens in parenteral drugs using the human whole blood pyrogen test, the rabbit in vivo pyrogen test and the LAL test. European Journal of Parenteral Science 5(2):39-44.

Levin J, Bang FB (1964). A description of cellular coagulation in the Limulus. Bull Johns Hopkins Hosp 155: 337-345.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, November 1989: 17-18.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' *in vitro* test for pyrogens. J Immunol Methods 274: 209-220.

Schindler S, Bristow A, Cartmell T, Hartung T and Fennrich S. (2003). Comparison of the reactivity of human and rabbit blood towards pyrogenic stimuli. ALTEX 20:59-63

Snedecor, G. W., Cochran, W. G. (1989). Statistical methods, 8th edition, Ames, US: The Iowa State University Press.

Spreitzer I, Fischer M, Hartzsch K, Lüderitz-Püschel U and Montag T. (2000). Comparative Study of Rabbit Pyrogen Test and Human whole Blood Assay on Human Serum Albumin. ALTEX 19 (Suppl. 1):73-75.

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

13 CATCH-UP VALIDATION: Human WB/IL-1 in vitro Pyrogen Test using 96-wells plates.

13.1 Rationale

Throughout the study described in the previous part of this BRD, tubes are employed as a container during the incubation of the blood sample. However using 96 wells plates as a container is obviously more convenient. Although a significant impact on the accuracy of the WB/IL-1 assay is not expected, the influence of 96-wells plate was studied in an additional catch-up validation study (trail plan attached in Appendix A), while applying the same study plan as in the main part of this BRD. This variant is indicated as the 96-wells WB/IL-1 method in the remainder of this section.

13.2 Test Method Protocol Components

The method follows the original standard protocol, with the obvious exemption of using 96-wells plates during incubation of the fresh blood (20 µl per well) with the samples of interest. Details of the test procedure are given under point 7: 7A - fresh blood using 96-well plates in test method protocol CRYO WB/IL-1: *Human Whole Blood Pyrogen Test in 96-well Plates Using Fresh or Cryopreserved Blood(electronic file name: SOP CRYO WB IL 1)* which was used during the catch-up validation (see Appendix A). The released IL-1□ is assessed using the standard IL-1□ ELISA.

13.3 Substances Used for Validation

The same 10 parenteral drugs used to determine sensitivity and specificity (see table 3.3.1.) were used for the catch-up validation. Again, each test item was tested after spiking at its individual MVD, thus came with their own specific set of 5 endotoxin spike solutions: 0.0, 0.25, 0.5, 0.5 and 1.0 EU/ml. The test items were assessed with 5 different endotoxin levels at 3 independent test facilities, yielding a total of 150 data points, biometrically considered to be sufficient for further analysis.
The same three drugs (table 3.3.2) as used for the prevalidation of the CRYO WB/IL-1 method (see BRD CRYO WB/IL-1) were employed. Each drug was tested at an interference free dilution and spiked with 0.0, 0.0, 0.5 and 1.0 EU/ml. The samples were tested at each of the 3 laboratories. The results were used to provide a preliminary estimate of the interlaboratory reproducibility and accuracy.

13.4 Preliminary estimate of the Test Method Accuracy

In short, three test substances were spiked with WHO-LPS at four levels (0, 0, 0.5 and 1.0 EU/ml respectively), which resulted in 12 samples with a balanced design with regard to positive and negative samples (i.e. samples expected to be pyrogenic and non-pyrogenic respectively. These 12 samples were tested in three laboratories (See figure 13.4.1)

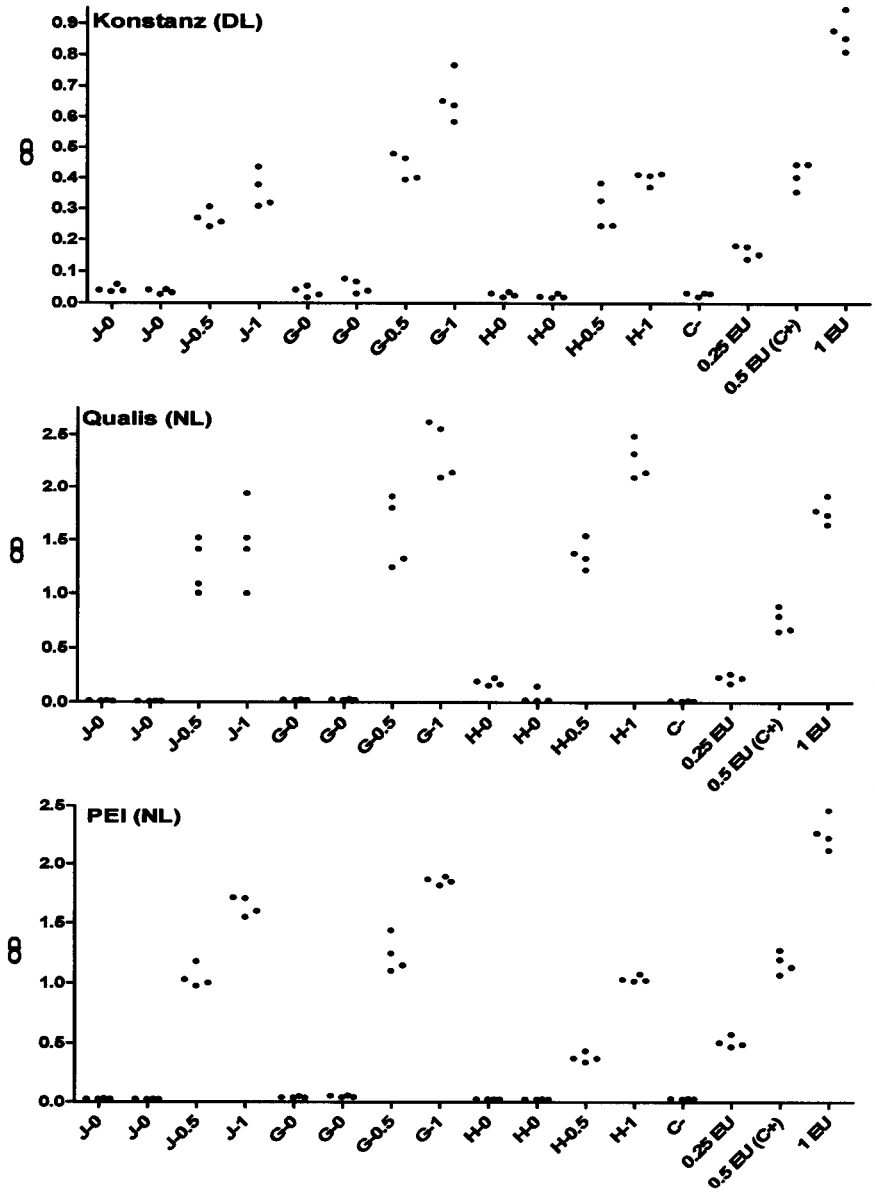


Figure 13.4.1: Prevalidation data for 96-wells WB/IL-1 of the three involved laboratories. The treatments and controls are abbreviated (indicating the endotoxin contamination in EU. (J = Jonosteril: G = Gelafundin; H = Haemate; C- = saline with 0 EU; C+ = positive control)

As Figure 13.4.1 only gives an indication about variability of replicates, the CVs were calculated for each treatment or control for all laboratories (Figure 13.4.2). In general the CVs were smaller than 30% and only two treatments produced a CV larger than 45%. Furthermore a tendency for larger CV of endotoxin-free samples/treatments was observed, as the background OD-level was lower compared to the equivalent assays, described in the main part of the BRD.

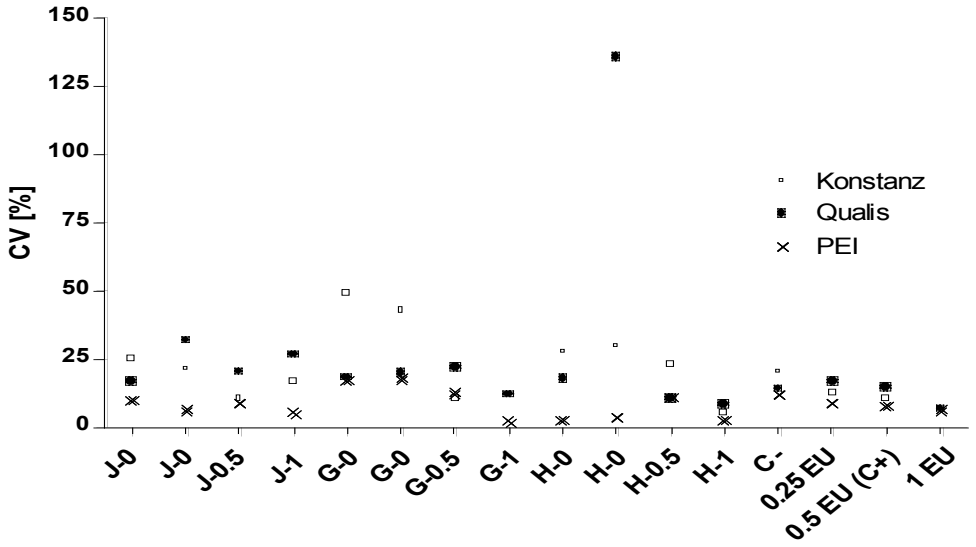


Figure 13.4.2: Coefficients of variation of the prevalidation data from 96-wells WB/IL-1 for the three involved laboratories. The treatments and controls are abbreviated indicating the endotoxin contamination in EU. (J = Jonosteril: G = Gelafundin; H = Haemate; C- = saline with 0 EU; C+ = positive control)

Application of the PM to these data resulted in the classifications summarized in Table 13.4.1 .Ten out of the twelve spikes were classified in the same way in all laboratories. Comparing the laboratories pair wise, showed that 32 of the total of 36 single comparison, i.e. 88.9%, resulted in the same classification.
Assessing in the final step preliminarily the predictive capacity, revealed that all negative samples were classified correctly and that two 0.5-EU spikes (Konstanz: J-0.5; PEI: H-0.5), which are at the rabbit classification threshold, were classified false negative.

BRD: WB/IL-1 March, 2006

Table 13.4.1: Classification by the 96-wells WB/IL-1 of the spikes in the prevalidation in the three involved laboratories.

drug	spike in EU	laboratory Konstanz	Qualis	PEI
Jonosteril	0	0	0	0
	0	0	0	0
	0.5	**0**	1	1
	1	1	1	1
Gelafundin	0	0	0	0
	0	0	0	0
	0.5	1	1	1
	1	1	1	1
Haemate	0	0	0	0
	0	0	0	0
	0.5	1	1	**0**
	1	1	1	1

Table 13.4.2: Preliminary estimate of interlaboratory reproducibility: Assessed by testing of 3 substances, spiked 4 times. One run of 12 samples by three different laboratories.

Laboratories	Interlaboratory Reproducibility	Number of equal predictions
DL - NL1	91.7%	11 / 12
DL - NL2	83.3%	10 / 12
NL1 – NL2	91.7%	11 / 12
Mean	88.9%	
same result in all laboratories	83.3%	10 / 12

DL =Konstanz; NL1 = Qualis; NL2 = PEI

A 2x2 contingency table was constructed (table 13.4.3), from which a preliminary estimates of sensitivity and specificity can easily be derived.

Table 13.4.3: 2x2 contingency table. The prediction model applied to a preliminary validation study with 96-wells WB/IL-1. Three different substances were assessed in three different laboratories (derived from table 13.4.1)

		True status of samples		**Total**
		+	-	
PM	+	16	0	16
	-	2	18	20
	Total	18	18	36

The specifications of specificity and sensitivity described in section 5.3 were applied to these results. The specificity (Sp) of the 96-wells WB/IL-1 assay is 100% and the sensitivity (Se) calculated for this data set is 88.9%. As outlined previously the specificity is overestimated and the sensitivity is underestimated as a result of the design of this part of the study.

Conclusion: Regarding the inherent variability of the assay method, the 96-wells WB/IL-1 showed good results. The result of the prevalidation show that the interlaboratory reproducibility and the predictive capacity in terms of specificity and sensitivity of the 96-wells WB/IL-1 are comparable with the WB/IL-1 using tubes.

13.5 Test Method Accuracy

To assess **accuracy** of the proposed test method 10 substances (listed in table 3.3.1), were spiked with five different concentrations of the WHO-LPS (one of which is negative). To permit the application of the chosen prediction model, each drug was diluted to its individual MVD, which has been calculated using the ELC to that drug (listed in table 3.3.1.) Each substance had their own specific set of endotoxin spike solution, ensuring that after spiking the undiluted drug, it contained 0.0; 0.25; 0.5; 0.5 and 1.0 EU/ml respectively when tested at its MVD. A detailed description of the sample preparation was supplied to the three independent laboratories (as shown in table 5.1.2). To put more weight to the result of this part of the validation, the spikes were blinded and coded by QA ECVAM. Accuracy was assessed by applying the PMl to the results and evaluating the concordance in a two by two table.

BRD: WB/IL-1 March, 2006

Table 13.5.1: Results of the validation study of 10 drugs, spiked with WHO-LPS at 0.0, 0.25, 0.5, 0.5 and 1.0 IU/ml, respectively and tested in 3 different laboratories. Samples and spikes were blinded. Classifications after applying the prediction model.

drug (code)	**spike EU/ml**	**"truth"**	**results** PEI	Qualis	Novartis
Beloc (BE)	0.00	0	0	0	0
	0.25	0	**1**	0	0
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Binotal (BI)	0.00	0	0	0	0
	0.25	0	0	**1**	0
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Ethanol 13% (ET)	0.00	0	0	0	CV
	0.25	0	0	**1**	0
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Fenistil (FE)	0.00	0	0	0	0
	0.25	0	CV	**1**	**1**
	0.50	1	CV	1	1
	0.50	1	CV	1	1
	1.00	1	1	1	1
Glucose 5% (GL)	0.00	0	0	0	0
	0.25	0	0	**1**	**1**
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
"Drug A"	0.00	0	CV	0	0
0.9% NaCl (LO)	0.25	0	0	0	0
	0.50	1	CV	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
MCP (ME)	0.00	0	0	0	0
	0.25	0	0	0	1
	0.50	1	**0**	1	1
	0.50	1	1	1	1
	1.00	1	CV	1	1

BRD: WB/IL-1 March, 2006

drug (code)	spike EU/ml	"truth"	results PEI	Qualis	Novartis
"Drug B"	0.00	0	0	0	0
0.9% NaCl (MO)	0.25	0	0	0	CV
	0.50	1	CV	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Orasthin (OR)	0.00	0	0	0	0
Syntocinon	0.25	0	0	0	0
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Sostril (SO)	0.00	0	0	0	0
	0.25	0	CV	0	**1**
	0.50	1	1	1	CV
	0.50	1	1	1	1
	1.00	1	1	1	1

"0"denotes "non-pyrogenic"; "1" denotes "pyrogenic".
CV = sample showed a variability resulting in exclusion, i.e. CV > 45% and no significant outlier present. nq = not qualified according to quality criteria, i.e. failure of PPCs and PCs
False classifications are in bold/colour type.

Of the 150 available data for the 96-wells WB/IL-1 method, eleven sets of 4 replicates showed a high variability resulting in exclusion, i.e. CV > 45% and no significant outliers present. Therefore 139 data in total could be used to estimate the specificity and sensitivity of the 96-wells WB/IL-1 method. The results are shown separately for each participating laboratory (table 13.5.2) as well as combined for all these laboratories (table 13.5.3).

The specificity that can be estimated from the available results for DL, NL1 and NL2 is 94.1%, 80% and 77.8% respectively The estimated sensitivity of the 96-wells WB/IL-1 assay was excellent of all three participating laboratories: 96%, 100% en 100% respectively (calculated from results in table 13.1.2).

Table 13.5.2: 2x2 contingency table. Prediction model applied to the 96-wells WB/IL-1 test result of 10 different substances assessed in three different laboratories. Results of each laboratory separately (DL, NL1 and NL2= PEI, Qualis and Novartis respectively).

Results DL		True status of samples +	-	Total
PM	+	24	1	25
	-	1	16	17
Total		25	17	42

Results NL1		**True status of samples** +	-	**Total**
PM	+	30	4	34
	-	0	16	16
Total		30	20	50

Results NL2		**True status of samples** +	-	**Total**
PM	+	29	4	33
	-	0	14	14
Total		29	18	47

The specificity of the combined results of the three laboratories of the 96-wells WB/IL-1 assay is 83.6% (46/(46+9)*100%), 95% confidence interval [0.712-0.922]. The sensitivity equals 98.8% (83/(83+1) *100%), 95% confidence interval [0.935-0.999]. (Summarized in table 13.5.3 and 13.5.4).

Table 13.5.3 2x2 contingency table. Prediction model applied to the 96-wells WB/IL-1 test result of 10 different substances assessed in three different laboratories. Combined results.

		True status of samples +	-	**Total**
PM	+	83	9	92
	-	1	46	47
	Total	84	55	139

Table 13.5.4 Specificity and sensitivity of the 96-wells WB/IL-1 assay

	N total	**N correctly identified**	**proportion**	**95% CI lower limit**	**95% CI upper limit**
Specificity (Sp)	55	46	83.6%	0.712	0.922
Sensitivity (Se)	84	83	98.8%	0.935	0.999

13.6 Test Method Reliability (Reproducibility)

The interlaboratory reproducibility of the 96-wells WB/IL-1 method was assessed from the results of the validation test with 10 substances spiked with 5 separate spikes. The reproducibility varied from 88.1% to 91.5% between two laboratories. Also the estimated reproducibility between the three participating laboratories was very satisfactory (84.6%)

Table 13.6.1.: Interlaboratory reproducibility, 96-wells WB/IL-1: Assessed by testing of 10 substances, spiked 5 times. One run of 50 samples by three different laboratories.

Laboratories	**Interlaboratory Reproducibility**	**Number of equal predictions**
DL - NL1	88.1%	37 / 42
DL - NL2	89.7%	35 / 39
NL1 – NL2	91.5%	43 / 47
Mean	89.8%	
same result in all laboratories	84.6%	33 / 39

13.7 Conclusion

In this catch-up validation study it is shown that the interlaboratory reproducibility of the WB/IL-1 assay could be improved with the 96-wells approach. All three laboratories found the same in 84.6% of the result, whereas only 57% was predicted the same with the original WB/IL-1. Also the predictive capacity of the 96 wells test in terms of specificity and sensitivity was very promising. In this catch-up validation study the specificity of the 96-wells was lower then for the tubes (83.6% versus 93.2%), but still very satisfactory. However, the sensitivity was considerably improved (98.8% versus 72.7%). It appears that implementation of the 96-wells plates is a remarkable improvement of the well established WB/IL-1 test.

14 Supporting Materials (Appendices)

14.1 Standard operating procedure (SOP) of the proposed method

Provide the complete, detailed protocol for the proposed test method.

Appendix A includes the test method protocol *WB/IL-1: Human Whole Blood Pyrogen (electronic file name: SOP-WB IL 1*) and the protocol used for the validation study *("Human Whole Blood Pyrogen Test - Standard Operating Procedure for the Validation Phase" marked with internal identifierSop-WBT-KNv02; electronic file name: SOP WB-IL-1 validation)*.
Regarding the WB/IL-1 using fresh blood and 96 well plates (Section 13 of this BRD), the method is described under point 7: 7A - fresh blood using 96-well plates in test method protocol CRYO WB/IL-1: *Human Whole Blood Pyrogen Test in 96-well Plates Using Fresh or Cryopreserved Blood(electronic file name: SOP CRYO WB IL 1)*.

The trial plan of the validation and catch-up validation study are also included in Appendix A.

14.2 Standard operating Procedure (SOP) of the reference method

Provide the detailed protocol(s) used to generate reference data for this submission and any protocols used to generate validation data that differ from the proposed protocol.

14.3 Publications

Provide copies of all relevant publications, including those containing data from the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method with established performance standards.

Not all of the publications cited in the BRD (see section 12) are included as hardcopies in Appendix B, e.g. publications on statistical methods are not given.

Remark: The same set of hardcopies was included into Appendix B of all of the 5 submitted BRDs. Therefore some of the publications in Appendix B might not be referenced in the current BRD nor included in the list of the publications in section 12. Three general publications were added, which are not cited in any BRD but might be useful as background information to the validation study: the ECVAM publications on validation (Balls et al, 1995; Curren et al, 1995; Hartung et al, 2004). Several publications were included, which either give more background information on the human fever reaction or report on specific studies using *in vitro* pyrogen tests.

Part 1:
List of hard copies

Andrade SS, Silveira RL, Schmidt CA, Junior LB, Dalmora SL. (2003) Comparative evaluation of the human whole blood and human peripheral blood monocyte tests for pyrogens. Int J Pharm. Oct 20;265(1-2):115-24.

Balls et al. (1995) Practical aspects of the validation of toxicity test procedures. ECVAM Workshop Report 5. ATLA 23, 129-147.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Carlin G, Viitanen E. (2003) In vitro pyrogenicity of a multivalent vaccine: Infanrix. PHARMEUROPA, 15(3), 418-423.

Curren et al (1995) The Role of prevalidation in the development, validation and acceptance of Alternative Methods. ECVAM Prevalidation Task Force Report 1. ATLA 23, 211-217.

De Groote, D., Zangerle, P.F., Gevaert, Y., Fassotte, M.F., Beguin, Y., Noizat-Pirenne, F., Pirenne, J., Gathy, R., Lopez, M., Dehart I., (1992). Direct stimulation of cytokines (IL-1 beta, TNF-alpha, IL-6, IL-2, IFN-gamma and GM-CSF) in whole blood. I. Comparison with isolated PBMC stimulation. Cytokine 4, 239.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Fennrich S, Wendel A and Hartung T. (1999). New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 16:146-149

Fennrich S, Fischer M, Hartung T, Lexa P, Montag-Lessing T, Sonntag H-G, Weigandt M und Wendel A. (1999). Detection of endotoxins and other pyrogens using human whole blood. Dev. Biol. Standards 101:131-139

Gaines Das et al (2004). Monocyte activation test for pro-inflammatory and pyrogenic contaminants of parenteral drugs: test design and data analysis. J. of Immunological Methods 288, 165-177.

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hartung et al (2004) A modular approach to the ECVAM principles on test validity. ATLA 32, 467-472.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005). Optimisation of pyrogen testing in parenterals according to different

pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Peterbauer A, Schindler S, et al (2005). International validation of novel pyrogen tests based on human monocytoid cells. J Immunol Methods 298: 161-173.

Jahnke M, Weigand, Sonntag H-G. (2000). Comparative testing for pyrogens in parenteral drugs using the human whole blood pyrogen test, the rabbit in vivo pyrogen test and the LAL test. European Journal of Parenteral Science 5(2):39-44

Nakagawa Y, Maeda H, Murai T. (2002) Evaluation of the in vitro pyrogen test system based on proinflammatory cytokine release from human monocytes: comparison with a human whole blood culture test system and with the rabbit pyrogen test. Clin Diagn Lab Immunol. May;9(3):588-97.

Pasare, C. and Medzhitov, R. (2003). Toll pathway-dependent blockade of CD4+CD25+ T cell-mediated suppression by dendritic cells. Science 299: 1033-1036.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1998) The detection of pyrogens in blood products using an ex vivo whole blood culture assay. J Immunoassay. May-Aug;19(2-3):95-111.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1999) Differentiation between endotoxin and non-endotoxin pyrogens in human albumin solutions using an ex vivo whole blood culture assay. J Immunoassay. Feb-May; 20(1-2):79-89.

Poole S, Thorpe R, Meager A, Hubbard AJ, Gearing AJ (1988). Detection of pyrogen by cytokine release. Lancet 8577, 130.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, November 1989: 17-18.

Poole S, Musset MV (1989). The International Standard for Endotoxin: evaluation in an international collaborative study. J Biol Stand 17: 161-171.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' in vitro test for pyrogens. J Immunol Methods 274: 209-220.

Ray A, Redhead K, Selkirk S, Polle S (1991) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211-218.

Reddi, K., Nair, S.P. et al (1996). Surface-associated material from the bacterium Actinobacillus actinomycetemcomitans contains a peptide which, in contrast to lipopolysaccharide, directly stimulates fibroblast interleukin-6 gene transcription. Eur. J. Biochem. 236: 871-876.

Schindler S, Bristow A, Cartmell T, Hartung T and Fennrich S. (2003). Comparison of the reactivity of human and rabbit blood towards pyrogenic stimuli. ALTEX 20:59-63.

Schindler et al (2004) Cryopreservation of human whole blood for pyrogenicity testing. Journal of Immunological Methods 294, 89–100

Schins RP, van Hartingsveldt B, Borm PJ. Ex vivo cytokine release from whole blood. A routine method for health effect screening. Exp Toxicol Pathol. 1996 Nov; 48(6):494-6.

Spreitzer I, Fischer M, hartzsch K, Lüderitz-Püschel U and Montag T. (2000). Comparative Study of Rabbit Pyrogen Test and Human whole Blood Assay on Human Serum Albumin. ALTEX 19 (Suppl. 1):73-75

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

Part 2:

List of Diploma theses, reports and/or PhDs etc. concerning the WB/IL-1 test (IPT: In vitro Pyrogen Test)

1. Final report for the BMBF (Bundesministerium für Bildung und Forschung) (University of Konstanz, 2000). „Evaluierung und Prävalidierung eines Vollblutmodelles zum Ersatz des Pyrogentests am Kaninchen (DAB10)“, Phase I, (“Evaluation and prevalidation of a whole blood assay for the replacement of the pyrogentest with rabbits”), July 1th, 1997 – June 30th, 2000, No. 0311424
2. Final report for the BMBF (Bundesministerium für Bildung und Forschung) (Langen, Paul-Ehrlich-Institut, 2000). „Evaluierung und Prävalidierung eines Vollblutmodelles zum Ersatz des Pyrogentests am Kaninchen (DAB10)“, Phase I, (“Evaluation and prevalidation of a whole blood assay for the replacement of the pyrogentest with rabbits”), July 1th, 1997 – June 30th, 2000, No. 0311425
3. PhD-Thesis from Markus Weigandt at the Ruprecht-Karls-University of Heidelberg, institute of hygiene (Director: Prof. H.-G. Sonntag): Der humane Vollblut-Pyrogentest: Optimierung, Validierung und Vergleich mit den Arzneibuchmethoden” (The human whole blood pyrogen test: optimization, validation and comparision with methods regulated in the pharmacopoeias), 2000
4. Master Thesis (Master of Science: MSc), Karin Kullmann: „Adaptation des *In vitro* Pyrogen Tests (IPT) für prothetische Materialien“ (“Adaptation of the *in vitro* pyrogen test (IPT) to medical devices”), Technical University of Furtwangen, July 2002
5. Final report for the BMBF (Bundesministerium für Bildung und Forschung) (Langen, Paul-Ehrlich-Institut, 2004). „Ersatz des Pyrogentests am Kaninchen durch einen Vollbluttest“, Phase II, (“replacement of the rabbit experiment with the whole blood test”), October 1th, 2000 – September 30th, 2003, No. 0311424A
6. Final report for the BMBF (Bundesministerium für Bildung und Forschung) (University of Konstanz, 2004). „Ersatz des Pyrogentests am Kaninchen durch einen Vollbluttest“, Phase II, (“replacement of the rabbit experiment with the whole blood test”), September 1th, 2000 – August 31th, 2003, No. 0311424A

7. Brazil/Germany Cooperation Project: final report for the BMBF (Bundesministerium für Bildung und Forschung). „Validation of *in vitro* Cytokine Release Assay (Whole Blood Assay) for Controlling the Quality of Human Injectable Products“ for bilaterial Cooperation in Science and Technology (Germany – Brazil), April 1th 2002-March 31th 2004, No. BRA 02/004
8. Cuba/Germany Cooperation Project: final report for the BMBF (Bundesministerium für Bildung und Forschung). „Pyrogenicity Testing by Human Whole Blood“ for bilaterial Cooperation in Science and Technology (Germany – Cuba), January 1th , 2001- December 31th, 2003, No. CUB 00/022
9. Final report for the BMWa (Bundesministerium für Wirtschaft und Arbeit): “Entwicklung einer humanrelevanten Messtechnik für luftgetragene Toxine mit humanem Vollblut“ (development of a human relevant measurement for air-borne toxins with human whole blood), Sept 3th 2001– Sept. 30th 2003, No. KF 0317101KRF1
10. Postdoctoral lecture qualification (Habilitation), Bert Zucker, “Luftgetragene Endotoxine in Tierställen“ (“air-borne pyrogens in a stable”), Institut für Tier- und Umwelthygiene an der freien Universität Berlin, Berlin, 2004
11. Manuscript for the DIF (Deutsches Industrieforum, DIF-Fachtagung), Stefan Fennrich: “Pyrogenverunreinigungen an medizinischen Oberflächen. *In vitro* pyrogen-Test (IPT) als humanrelevantes Prüfverfahren“ (Contamination with pyrogens on medical surfaces: the *in vitro* pyrogen test (IPT) as a human specific method), Würzburg, June 21th -22 th, 2004, No. DIF 21/78/FE

Part 3: Further publications concerning the WB/IL-1 test (IPT)

1. Hartung T und Wendel A. Die Erfassung von Pyrogenen in einem humanen Vollblutmodell. ALTEX 1995,12:70-75
2. Fennrich S, Fischer M, Hartung T, Lexa P, Montag-Lessing T, Sonntag H-G, Weigandt M und Wendel A. Entwicklung und Evaluierung eines Pyrogentests mit menschlichem Blut. ALTEX 1998, 15:123-128
3. Fennrich S, Berthold S, Weigandt M, Lexa P, Sonntag H-G, Hartung T, Wendel A. Tagungsberichte, Pyrogentestung mit humanem Blut. Der Tierschutzbeauftragte 2, 1999, 102-107
4. Bonenberger J, Diekmann W, Fennrich S, Fischer M, Friedrich A, Hansper M, Hartung T, Jahnke M, Löwer J, Montag T, Petri E, Sonntag H-G, Weigand M, Wendel A, Zucker B. Pyrogentestung mit Vollblut. Zusammenfassung eines Status Workshops am Paul-Ehrlich-Institut, Langen, am 22.11.1999. Springer Verlag, Bundesgesundheitsbl-Gesundheitsforsch-Gesundheitsschutz, 2000, 43:525-533
5. Petri E, van de Ploeg A, Habermaier B und Fennrich S. Improved detection of pyrogenic substances on polymer surfaces with an ex vivo human whole-blood assay in comparison to the Limulus amoebocyte lysate test. In: Progress in the Reduction, Refinement and Replacement of Animal Experimentation. Editors: Balls M, van Zeller A-M, Halder M.E., Elsevier Science, 2000, 339-345

6. Hartung T, Fennrich S, Fischer M, Montag-Lessing T und Wendel A. Prevalidation of an Alternative to the rabbit test based on human whole blood. In: Progress in the Reduction, Refinement and Replacement of Animal Experimentation. Editors: Balls M, van Zeller A-M, Halder M.E., Elsevier Science, 2000, 991-999
7. Fennrich S, Zucker Bert and Hartung T. Beispiel eines neuen Einsatzbereichs des humanen Vollbluttests: Entwicklung eines Messverfahrens zur Abschätzung der gesundheitlichen Gefährdung durch luftgetragene mikrobielle Verunreinigungen. ALTEX 2001, 18:41-46
8. Thomas Hartung, Ingeborg Aaberge, Susanne Berthold, Gunnar Carlin, Emmanuelle Charton, Sandra Coecke, Stefan Fennrich, Matthias Fischer, Martin Gommer, Marlies Halder, Kaare Haslov, Michael Jahnke, Thomas Montag-Lessing, Stephen Poole, Leonard Schechtman, Albrecht Wendel and Gabriele Werner-Felmayer. Novel Pyrogen Tests Based on the Human Fever Reaction, The report and Recommendations of ECVAM Workshop 43, 2001, ATLA 29, 99-123
9. Fennrich S, Atemluft, gesund oder gefährlich…..das ist hier die Frage! Tagungsberichte. ALTEX 2002, 19: 43-45
10. Hartung T. Comparison and validation of novel pyrogen tests based on the human fever reaction. ATLA 2002, 30 (Suppl. 2):49-51
11. Morath S, Stadelmaier A, Geyer A, Schmidt RR and Hartung T. Synthetic lipoteichoic acid from Staphylococcus aureus is a potent stimulus of cytokine release. J. Exp. Med., 2002, 195:1635-1640
12. Morath S, Geyer A, Spreitzer I, Hermann C and Hartung T. Structural decomposition and heterogeneity of commercial lipoteichoic acid preparation. Infect. Immun. 2002, 70:938-944
13. Kindinger I, Fennrich S, Zucker B, Linsel G and Hartung T. Determination of air-borne pyrogens by the *in vitro* pyrogen test (IPT) based on human whole blood cytokine response. VDI-Bericht 1656 2002, 499-507
14. Schindler S, Reichstein S, Kindinger I, Hartung T, Fennrich S. New Ways in Pyrogen Testing: Replacing the Rabbit Experiment. Screening, Trends in Drug Discovery May, GIT Verlag, 2-3/2003, 4: 51-53
15. Zucker B A, Linsel G, Fennrich S, Müller W. Die Charakterisierung der entzündungsauslösenden Potenz von Bioaerosolen mittels Interleukinfreisetzung aus humanem Vollblut. Springer, VDI-Verlag. Gefahrstoffe Reinhaltung der Luft (Air Quality Control) 4, 2004, 155-158

14.4 Original data

Include all available non-transformed original data for both the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method with established performance standards.

NOTE: The original data of the ELISA-plate reader were collected by S.Hoffman and ECVAM. These are available on the CD, which goes with the BRD.

BRD: WB/IL-1 March, 2006

14.5 Performance standards

If appropriate performance standards for the proposed test method do not exist, performance standards for consideration by NICEATM and ICCVAM may be proposed. Examples of established performance standards can be located on the ICCVAM / NICEATM web site at http://iccvam.niehs.nih.gov.

BRD: WB/IL-1 March, 2006

APPENDIX A

Trial plan "Comparison And Validation Of Novel Pyrogen Tests Based On The Human Fever Reaction" Acronym: Human (e) Pyrogen Test

Detailed protocol WB/IL-1: *"Human whole blood pyrogen test"(electronic file name: SOP WB-IL-1)*

Detailed protocol WB/IL-1: *Human Whole Blood Pyrogen Test - Standard Operating Procedure for the Validation Phase" marked with internal identifierSop-WBT-KNv02 (electronic file name: SOP WB-IL-1 validation)*

Trial plan "Catch-up Validation of Novel Pyrogen Tests Based on the Human Fever Reaction"

Detailed protocol 96-wells WB/IL-1: Method 7A in *Human Whole Blood Pyrogen Test in 96-well Plates Using Fresh or Cryopreserved Blood(electronic file name: SOP CRYO WB IL 1*).

BRD: WB/IL-1 March, 2006

APPENDIX B

Not all of the publications cited in the BRD (see section 12) are included as hardcopies in Appendix B, e.g. publications on statistical methods are not given.

Remark: The same set of hardcopies was included into Appendix B of all of the 5 submitted BRDs. Therefore some of the publications in Appendix B might not be referenced in the current BRD nor included in the list of the publications in section 12. Three general publications were added, which are not cited in any BRD but might be useful as background information to the validation study: the ECVAM publications on validation (Balls et al, 1995; Curren et al, 1995; Hartung et al, 2004). Several publications were included, which either give more background information on the human fever reaction or report on specific studies using *in vitro* pyrogen tests.

List of hard copies

Andrade SS, Silveira RL, Schmidt CA, Junior LB, Dalmora SL. (2003) Comparative evaluation of the human whole blood and human peripheral blood monocyte tests for pyrogens. Int J Pharm. Oct 20;265(1-2):115-24.

Balls et al. (1995) Practical aspects of the validation of toxicity test procedures. ECVAM Workshop Report 5. ATLA 23, 129-147.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Carlin G, Viitanen E. (2003) In vitro pyrogenicity of a multivalent vaccine: Infanrix. PHARMEUROPA, 15(3), 418-423.

Curren et al (1995) The Role of prevalidation in the development, validation and acceptance of Alternative Methods. ECVAM Prevalidation Task Force Report 1. ATLA 23, 211-217.

De Groote, D., Zangerle, P.F., Gevaert, Y., Fassotte, M.F., Beguin, Y., Noizat-Pirenne, F., Pirenne, J., Gathy, R., Lopez, M., Dehart I., (1992). Direct stimulation of cytokines (IL-1 beta, TNF-alpha, IL-6, IL-2, IFN-gamma and GM-CSF) in whole blood. I. Comparison with isolated PBMC stimulation. Cytokine 4, 239.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Fennrich S, Wendel A and Hartung T. (1999). New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 16:146-149

Fennrich S, Fischer M, Hartung T, Lexa P, Montag-Lessing T, Sonntag H-G, Weigandt M und Wendel A. (1999). Detection of endotoxins and other pyrogens using human whole blood. Dev. Biol. Standards 101:131-139

Gaines Das et al (2004). Monocyte activation test for pro-inflammatory and pyrogenic contaminants of parenteral drugs: test design and data analysis. J. of Immunological Methods 288, 165-177.

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hartung et al (2004) A modular approach to the ECVAM principles on test validity. ATLA 32, 467-472.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005). Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Peterbauer A, Schindler S, et al (2005). International validation of novel pyrogen tests based on human monocytoid cells. J Immunol Methods 298: 161-173.

Jahnke M, Weigand, Sonntag H-G. (2000). Comparative testing for pyrogens in parenteral drugs using the human whole blood pyrogen test, the rabbit in vivo pyrogen test and the LAL test. European Journal of Parenteral Science 5(2):39-44

Nakagawa Y, Maeda H, Murai T. (2002) Evaluation of the in vitro pyrogen test system based on proinflammatory cytokine release from human monocytes: comparison with a human whole blood culture test system and with the rabbit pyrogen test. Clin Diagn Lab Immunol. May;9(3):588-97.

Pasare, C. and Medzhitov, R. (2003). Toll pathway-dependent blockade of CD4+CD25+ T cell-mediated suppression by dendritic cells. Science 299: 1033-1036.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1998) The detection of pyrogens in blood products using an ex vivo whole blood culture assay. J Immunoassay. May-Aug;19(2-3):95-111.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1999) Differentiation between endotoxin and non-endotoxin pyrogens in human albumin solutions using an ex vivo whole blood culture assay. J Immunoassay. Feb-May; 20(1-2):79-89.

Poole S, Thorpe R, Meager A, Hubbard AJ, Gearing AJ (1988). Detection of pyrogen by cytokine release. Lancet 8577, 130.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, November 1989: 17-18.

Poole S, Musset MV (1989). The International Standard for Endotoxin: evaluation in an international collaborative study. J Biol Stand 17: 161-171.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' in vitro test for pyrogens. J Immunol Methods 274: 209-220.

Ray A, Redhead K, Selkirk S, Polle S (1991) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211-218.

Reddi, K., Nair, S.P. et al (1996). Surface-associated material from the bacterium Actinobacillus actinomycetemcomitans contains a peptide which, in contrast to lipopolysaccharide, directly stimulates fibroblast interleukin-6 gene transcription. Eur. J. Biochem. 236: 871-876.

Schindler S, Bristow A, Cartmell T, Hartung T and Fennrich S. (2003). Comparison of the reactivity of human and rabbit blood towards pyrogenic stimuli. ALTEX 20:59-63.

Schindler et al (2004) Cryopreservation of human whole blood for pyrogenicity testing. Journal of Immunological Methods 294, 89–100

Schins RP, van Hartingsveldt B, Borm PJ. Ex vivo cytokine release from whole blood. A routine method for health effect screening. Exp Toxicol Pathol. 1996 Nov; 48(6):494-6.

Spreitzer I, Fischer M, hartzsch K, Lüderitz-Püschel U and Montag T. (2000). Comparative Study of Rabbit Pyrogen Test and Human whole Blood Assay on Human Serum Albumin. ALTEX 19 (Suppl. 1):73-75

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

BRD: WB/IL-1 March, 2006

APPENDIX C

List of abbreviations and definitions

Accuracy	The ability of a test system to provide a test result close to the accepted reference value for a defined property.
BET	The bacterial endotoxin test is used to detect or quantify endotoxins of gram-negative bacterial origin using amoebycte lysate from horseshoe crab (*Limulus polyphemus* or *Tachypleus tridentatus*
BRD	Background Review Document
CRYO WB/IL-1	Whole blood assay (using cryopreserved blood) with IL-1 as endpoint
CV	coefficient of variation
DL	Developing laboratory = laboratory which developed the method or the most experienced laboratory
ELC	Endotoxin limit concentration; maximum quantity of endotoxin allowed in given parenterals according to European Pharmacopoeia
Endotoxins	Endotoxins are a group of chemically similar cell-wall structures of Gram-negative bacteria, i.e. lipopolysaccharides
ELISA	Enzyme linked immunosorbent assay
EU/ml	European Units per ml
IL-1	interleukin 1
IL-6	interleukin 6
Intralaboratory reproducibility	A determination of the extent that qualified people within the same laboratory can independently and successfully replicate results using a specific protocol at different times.
Interlaboratory reproducibility	A measure of the extent to which different qualified laboratories, using the same protocol and testing the same substances, can produce qualitatively and quantitatively similar results. Inter-laboratory reproducibility is also referred to as between-laboratory reproducibility.
KN	University of Konstanz (Konstanz, Germany), developing laboratory WB/IL-1 and CRYO WB/IL-1
LPS	lipopolysaccharides
MM6	MONO MAC-6 cell line
MM6/IL-6	In vitro pyrogen test using MM6 cell line and IL-6 release

	as an endpoint
MVD	Maximum valid dilution; the MVD is the quotient of the ELC and the detection limit
NIBSC	National Institute for Biological Standards and Control (London, UK), developing laboratory for WB/IL-6
NL	naïve laboratory = laboratory with non or minor experience with the method
NPC	negative product control (clean, released lot of the nominated product under test)
Novartis	Novartis (Basel, Switzerland), developing laboratory PBMC/IL-6
OD	optical density
PBMC	Peripheral blood mononuclear cells
PBMC/IL-6	In vitro pyrogen test using fresh peripheral blood mononuclear cells and IL-6 release as endpoint
PBMC-CRYO/IL-6	In vitro pyrogen test using cryopreserved peripheral blood mononuclear cells and IL-6 release as endpoint
PEI	Paul-Ehrlich Institut (Langen, Germany), participating laboratory
PM	prediction model = is an explicit decision-making rule for converting the results of the in vitro method into a prediction of in vivo hazard
PPC	positive product control (product under test spiked with 0.5 EU/ml of WHO-LPS (code 94/580)
Prevalidation study	A prevalidation study is a small-scale inter-laboratory study, carried out to ensure that the protocol of a test method is sufficiently optimised and standardised for inclusion in a formal validation study. According to the ECVAM principles, the prevalidation study is divided into three phases: protocol refinement, protocol transfer and protocol performance (Curren et al, ATLA 23, 211-217).
Pyrogens	fever-causing materials
Pyrogens, endogenous	endogenous pyrogens are messenger substances released by blood cells reacting to pyrogenic materials; e.g. IL-1, IL-6, TNF-α, prostaglandin E_2
Pyrogens, exogenous	exogenous pyrogens derive from bacteria, viruses, fungi or from the host himself
Reliability	Measures of the extent to which a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and interlaboratory

	reproducibility and intra-laboratory repeatability.
Relevance	Relevance of a test method describes whether it is meaningful and useful for a particular purpose. It is the extent to which the measurement result and uncertainty can accurately be interpreted as reflecting or predicting the biological effect of interest.
Repeatibility	Repeatability describes the closeness of agreement between test results obtained within a single laboratory when the procedure is performed independently under repeatability conditions, i.e. in a set of conditions including the same measurement procedure, same operator, same measuring system, same operating conditions and same location, and replicated measurements over a short period of time.
RIVM	National Institute of Public Health and the Environment (Bilthoven, The Netherlands), developing laboratory MM6/IL-6 method
Sensitivity	Sensitivity is the proportion of all positive/active substances that are correctly classified by a test method.
Specificity	Specificity is proportion of all negative/inactive substances that are correctly classified by a test method.
TMB	chromogenic substrate 3,3´,5,5´ -tetramethylbenzidine
TNF-α	tumour necrosis factor-α
USP	US Pharmacopoeia
Validation	Validation is the process by which the reliability and relevance of a procedure are established for a specific purpose
Validation study	A validation study is a large-scale interlaboratory study, designed to assess the reliability and relevance of an optimised method for a particular purpose
WB/IL-1	Whole blood assay (using fresh blood) with IL-1 release as endpoint
WB/IL-6	Whole blood assay (using fresh blood) with IL-6 release as endpoint
WHO	World Health Organization

[This Page Intentionally Left Blank]

Validation of Biomedical Testing Methods

Human Whole Blood Pyrogen Test

Standard Operating Procedure

Only the responsible of the GLP/QA Unit is allowed to make copies of this document. Extra examples can be obtained at the GLP/QA Unit. Quality Documents are only valid if they are signed by the responsible of the GLP/QA Unit in and provided with a blue copy number	Copy number

Identity

JOINT RESEARCH CENTRE
EUROPEAN COMMISSION
Institute for Health and Consumer Protection
ECVAM

Version number: _____
Applicable from: _____
Expired at: _____

Drafted by:	*Name* *Date* *Signature*	*Thomas Hartung,* *28/02/02*	*M. D. Ph.D.*
Reviewed by:	*Name* *Date* *Signature*		
Approved by:	*Name* *Date* *Signature*		
Issued by:	*Name* *Date* *Signature*		

*Owner/Trainer: Signature: Date:

PAGE OF CHANGES

Date of change/ Date of draft:	Version-number:	Changed page(s):	Summary of the change(s):	Changed by/Sign.:
01.04.01				
29.06.01		10		
20.09.01		2, 4, 5, 6, 8, 10, 11, 12, 13, 14 16		
14.12.01		2, 6, 8, 9, 10, 11		
23.02.02		2, 5, 6,7, 8, 9, 10,11		
28.2.02		2, 6, 7, 10, 12		

3

TABLE OF CONTENTS Page no.

1. INTRODUCTION 4
2. PURPOSE 5
3. SCOPE / LIMITATIONS 5
4. METHOD OUTLINE 5
5. DEFINITIONS / ABBREVIATIONS 6
6. MATERIALS 7
7. METHODS 9
8. DATA ANALYSIS AND ASSOCIATED ERRORS 12
9. PREDICTION MODEL 13
10. HEALTH SAFETY AND ENVIRONMENT 13
11. ANNEX 14
12. REFERENCES 16

1. INTRODUCTION

The whole blood pyrogen test (in vitro pyrogen test IPT) is a two-part assay for the detection of pyrogenic contamination. It involves incubation of the sample with human blood, followed by an enzyme immunoassay for the measurement of IL-1b.
A pyrogen is a substance that causes fever. Bacterial contaminations, which contain exogenous pyrogens, can be deadly. This problem is of great significance for drug safety.
Also, medical devices and biologically produced substances obtained from bacteria and other microorganisms may cause release of endogenous pyrogens (e.g., IL-1b).
Exogenous pyrogens include metabolic substances and cell-wall components of microorganisms. These substances are present during the "normal" course of an infectious disease. Infections by gram-negative and gram-positive bacteria are equal in frequency. Both of these bacterial types can activate the release of endogenous pyrogens, which cause fever through the thermoregulatory center in the brain. Although these reactions can occur during the "normal" course of an infectious disease, a deadly shock syndrome can occur in the worst case.
Due to these risks, product safety legislation demands rigorous quality checks for pyrogenic contamination of drugs and devices intended for parenteral use. For example, testing in rabbits for medical end products is required in Germany. Products in development and a few end products are allowed to be controlled by the Limulus assay. The first pyrogen assay, based on human whole blood stimulation by pyrogens,

Institute for Health and Consumer Protection

was developed by Hartung et al. (3,4).

2. PURPOSE

This assay simulates *in vitro* the normal human reaction to exogenous pyrogens. A few drops of human blood are mixed with the sample, and exogenous pyrogens in the sample are recognized by immunocompetent cells in the human blood.
These cells release IL-1b, which is measured by an integrated ELISA system.

3. SCOPE / LIMITATIONS

Limit of detection is ≤ 0,25 EEU/ml, not suitable for test samples interfering with blood cytokine release (see 8: Data analysis and associated errors).

4. METHOD OUTLINE

The procedure has two parts:
1) Incubation of the sample with (diluted) human blood
2) An enzyme immunoassay for the measurement of IL-1b.

Ad 1) Blood incubation
Diluted human whole blood is incubated for 10-24 hours together with saline and the sample in pyrogen-free reaction tubes. It is then centrifuged and the supernatant is taken off for further examination.

Ad 2) Capture of Endogenous Pyrogens (ELISA procedure)
Samples (supernatants of blood stimulation) are distributed into the wells of a microplate which are coated with monoclonal antibodies specific for IL-1b.

An enzyme-conjugated polyclonal antibody against IL-1b is added. During a 90-minute incubation, a sandwich complex consisting of two antibodies and the IL-1b is formed. Unbound material is removed by a wash step.

A chromogenic substrate (3,3´,5,5´ -tetramethylbenzidine, TMB) reactive with the enzyme label is added. Color development is terminated by adding a stop solution after 30 minutes. The resulting color, read at 450 nm, is directly related to the IL-1b concentration. Bi-chromatic measurement with a 600-690 nm reference filter is recommended.

5

Institute for Health and Consumer Protection

5. DEFINITIONS / ABBREVIATIONS

The following abbreviations are used in this work-book.

Ab	antibody
°C	degrees Celsius (Centigrade)
EC	endotoxin control
EEU	endotoxin equivalent unit
ELISA	Enzyme-Linked ImmunoSorbent Assay
ESS	Endotoxin Stabilizing Solution
EU	endotoxin unit of the international standard
h	hour
H_2SO_4	sulphuric acid
IL	interleukin
LPS	lipopolysaccharide (exogenous pyrogen from Gram-negative bacteria)
LTA	lipoteichoic acid (exogenous pyrogen from Gram-positive bacteria)
l	litre
µg	microgram
µl	microlitre
mg	milligram
ml	millilitre
min	minute
MAb	monoclonal antibody
NaCl	sodium chloride, 0,9%
nm	nanometre
PPC	positive product control
OD	optical density
rpm	rounds per minute
RT	room temperature
TMB	3,3´,5,5´-Tetramethylbenzidine
WDB	wash/dilution buffer
x g	x gravity

6

6. MATERIALS

6.1. Materials required and not provided

The components listed below are recommended, but equivalent devices may also be used: it is the user's responsibility to validate the equivalence.
For all steps excluding the ELISA procedure sterile and pyrogen-free materials have to be used (e.g. tips, containers, solutions).

6.1.1. Materials for Blood Incubation

A. Tube method

Equipment
· Incubator or thermoblock (37°C ± 1°C)
· Multipette
· Centrifuge (recommended)
· Vortex mixer

Consumables
· Sterile and pyrogen-free tips 100 µl and 1000 µl
· Heparinized tubes for blood sampling (e.g. Sarstedt S-MONOVETTE 7.5 ml, 15 IU/ml Li-Heparin)
· Sarstedt multifly needle set, pyrogenfree, for S-Monovette
· 1.5 ml closable, pyrogen-free reaction tubes
· Reservoir for saline
· Combitips for multipette, 10 ml and 2,5 ml for pipetting saline and blood
· Non-pyrogenic borosilicate test tubes or other qualified materials that can be used for the preparation of standards and for the dilution of samples.

B. Microtiter plate method

Equipment
·Incubator or thermoblock (37°C)
·Multipette or adjustable 20 to 100 µl pipetters
·Vortex mixer

Consumables
·Sterile and pyrogen-free tips 20 µl and 100 µl or
·Combitips for multipette, 2,5 and 1,0 ml
·Heparinized tubes for blood sampling(e.g. Sarstedt S-MONOVETTE 7.5 ml, 15 IU/ml Li-Heparin)
7

Institute for Health and Consumer Protection

·Sarstedt multifly needle set, pyrogen-free, for S-Monovette
·Non-pyrogenic tissue culture microtiter plate
·Reservoir for saline
·Non-pxrogenic borosilicate test tubes or other qualified materials that can be used for preparing standards and diluting samples

6.1.2. Materials for ELISA procedure

Equipment
· Multichannel pipettor
· Microplate mixer
· Microplate washer
· Microplate reader capable of readings at 450 nm (optional reference filter in the range of 600-690 nm)
· A software package for facilitating data generation, analysis, reporting, and quality control

Consumables
· Graduated cylinder and plastic storage container for Buffered Wash Solution

6.2. Materials Supplied in ELISA kit

Components supplied in that kit are *not* interchangeable with other lots of the same components.
IL-1b Ab-coated Microplate: One 96-well polystyrene microplate, packaged in a zip-lock foil bag, with desiccant. The plate consists of twelve strips mounted in a frame. Each strip includes eight anti-IL-1b ab-coated wells. Additionally, individual wells can be separated from the strip to enable the complete use of all the wells of a kit. Well positions are indexed by a system of letters and numbers (A through H, 1 through 12) embossed on the left and top edges of the frame. Store refrigerated: stable at 2-8°C until the expiration date marked on the label.
Enzyme-Labeled Antibody : One amber vial containing 16 ml of liquid reagent, ready-to-use. The reagent contains horseradish peroxidase-labeled, affinity-purified, polyclonal (rabbit) anti-IL-1b antibodies, with preservative. Store refrigerated: stable at 2-8°C for 30 days after opening, or until the expiration date marked on the label. Mix thoroughly before use. *Do not freeze.*
Endotoxin Stabilizing Solution: for reconstitution and dilution of the endotoxin control.
Endotoxin Control: One vial of an endotoxin control in a buffer matrix, with preservative. The control is supplied lyophilized. Store refrigerated: stable at 2-8°C until the expiration date marked on the label. At least 30 minutes before use, reconstitute control vial with saline. Prepare serial dilutions in ESS (see 7. Methods).
8

Institute for Health and Consumer Protection

Mix by vortexing. After preparation, the stock solution can be stored for up to 1 week hours at 4 °C.
LTA control: One vial of LTA control. The control is supplied lyophilized. Store refrigerated at 2-8°C until the expiration date. Before use, reconstitute the lyophilisate with 1 ml saline. Mix for at least a minute by vortexing. After preparation, the solution can be stored for up to 4 weeks at 4°C.
Saline: Three glass vials, each containing pyrogen-free saline. This is intended for the dilution of donor blood samples. Store refrigerated: stable at 2-8°C until the expiration date marked on the label. Use immediately after opening and discard unused volumes.
TMB/Substrate Solution: Two amber vials, each containing 11 ml of a buffered reagent, ready-to-use. The reagent contains a hydrogen peroxide substrate and 3,3´,5,5´-tetramethylbenzidine (TMB). Store refrigerated and protected from light: stable at 2-8°C until the expiration date marked on the label. *Do not freez*e.
Buffered Wash Solution Concentrate: One vial containing 75 ml of a concentrated (10X) buffered saline solution, with surfactants and preservative. Using a transfer container, dilute the contents of the vial with **675 ml** distilled or deionized water for a total volume of **750 ml**.
Store refrigerated: stable at 2-8°C for 30 days after preparation, or until the expiration date marked on the label.
For longer storage aliquot and freeze: stable at -20°C for 6 months.
Stop Solution: One vial containing an acidic solution, for terminating the color reaction. The reagent is supplied ready-to-use. Handle with care, using safety gloves and eye protection. Store refrigerated: stable at 2-8°C for 8 weeks after opening, or until the expiration date marked on the label.
Adhesive Microplate Covers: Two clear plastic adhesive covers. Remove backing and place over the top of the microplate during incubation to avoid evaporation.

7. METHODS

7.1. Blood Incubation

Blood Collection
Collect blood by venipuncture into heparinized tubes. The blood collection system must be pyrogen-free. The procedure calls for 100 µl of heparinized whole blood per assay. The blood can be stored in the collection tube at room temperature (15-28°C) for 4 hours. Incubation of the sample should be started within this time.

Note:

1 Blood donors should show no evidence of disease or need of medication during the last two weeks.
9

2 Each assay should include the Endotoxin Controls in duplicate (EC 0.5 in triplicate) and the saline control in triplicate.
3 Use disposable tip pipets to avoid contamination of reagents and samples.
4 During ELISA procedure, the wells should be washed carefully.
5 The test samples should be done in triplicate.
· The contents of the wells must be decanted or aspirated completely before pipetting wash solution.
· The wells should be covered during the incubation to avoid evaporation.
6 Deviations from the procedure (incubation time/temperature) may cause erroneous results. The ELISA procedure should be run without interruption. Diluted samples should be tested within an hour.

Endotoxin dilution

NOTE:
Quantitative IPT assays may use endotoxin concentrations of 5.0, 2.5, 1.0, 0.5 and 0.25 EU/ml + saline control in triplicate.

Qualitative IPT assays (threshold assays) should use the 0.5 EU/ml + saline control in triplicate.

Dissolve the contents of the vial with ESS according to directions stated in the Certificate of Analysis, yielding a stock solution = solution S

Solution	amount added to ESS	Volume of ESS	Resulting solution for use in blood incubation
Stock (5IU/ml)	500 µl	500 µl	Endotoxin Control (2,5 EEU/ml)
Endotoxin Control (2,5 EEU/ml)	400 µl	600µl	Endotoxin Control (1,0 EEU/ml)
Endotoxin Control (1,0 EEU/ml)	500 µl	500 µl	Endotoxin Control (0.5 EEU/ml)
Endotoxin Control (0.5 EEU/ml)	500 µl	500 µl	Endotoxin Control (0.25 EEU/ml)

LTA dilution

Reconstitute the vial with 1 ml saline. Mix by vortexing for 3 minutes.

10

Whole Blood Stimulation-test tube method

Perform incubation of blood samples in 1.5 ml pyrogen-free reaction tubes. Preferably, use a laminar-flow bench**.** All consumables and solutions have to be sterile and pyrogen-free.

With some substances, interference with the ELISA may occur. Therefore, it might be necessary to test the samples in different dilutions.

Step 1: add **1000 µl** saline into each reaction tube.

Step 2: add **100 µl** of each sample into the prepared reaction tubes or **100 µl** of the Endotoxin Control in duplicate (EC 0.5 in triplicate) and the negative control (saline) in triplicate.

Step 3: add **100 µl** of donor blood, mixed by gentle inversion, into each reaction tube.

Step 4: Close the tubes and invert them once or twice before starting the incubation.

Step 5: Incubate the closed reaction tubes in an incubator or a heating block overnight (10-24 hours) at 37°C ± 1°C.

Step 6: Mix the incubation tubes thoroughly by inverting the tubes. Incubations are to be centrifuged for 2 minutes at 10.000 g and the clear supernatant is used for the ELISA procedure. Take aliquots of ≥ **150 µl.**

The supernatants can be tested immediately by the ELISA System or may be stored at -20°C for testing at a later time.
Freeze additional aliquots.

Interference testing (PPC)

For each new sample, to determine whether it requires dilution prior to assay, perform the following experiment in triplicate.
The experiment checks for interference between the sample and the whole blood, and is needed only when the interference status of the sample has not yet been established. First assay 100 µl of the sample, undiluted, in combination with saline, Endotoxin Control (0.5 EEU) and whole blood, as follows

Step 1: add **900 µl** of saline into each tube

Step 2: add **100 µl** of (diluted) sample

Step 3: add **100 µl** of 1,0 EU/ml Endotoxin control
11

Step 4: add **100 µl** of donor blood, mixed by gentle inversion

Continue with Step 4 of Whole Blood Stimulation- test tube method procedure.

Whole blood stimulation- microtiter plate method

Step 1: Using a non-pyrogenic tissue culture treated microtiter plate, draw up an incubation plan designating the layout of endotoxin controls (i.g. 4x3), negative saline controls (3x), Gram-positive control (3x) and your samples (3x) in your assay (corresponding to template)

Step 2: Pipet 200 µl negative saline control into each of the reaction wells that will be used for the standards (endotoxin and Gram-positive control) and samples.

Step 3: Add 20 µl of endotoxin controls, Gram-positive control, negative saline control or samples into their respective reaction wells according to the prepared incubation plan.

Step 4: Add 20 µl of whole blood to all reaction wells.

Step 5: Cover with the dedicated plastic plate cover and mix thoroughly on a microtiter plate mixer.

Step 6: Transfer the mixed microtiter plate to a 37°C incubator for an overnight incubation (10 to 24 hours).

Step 7: Following the overnight incubation, remove plate from incubator place onto a plate mixer. Mix until all bloods cells have been re-suspended.

Step 8: The re-suspended blood mixtures may be ELISA tested for IL-1ß immediately or stored frozen at –20°C for testing at a later time (at least 150 µl).

Interference testing (PPC)

For each new sample, to determine whether it requires dilution prior to assay, perform the following experiment in triplicate.
The experiment checks for interference between the sample and the whole blood and is needed only when the interference status of the sample has not yet been established. First assay 20 µl of the sample, undiluted, in combination with saline, Endtoxin Control (1,0 EU/ml) and whole blood, as follows

12

Institute for Health and Consumer Protection

Step 1: Pipet **180 µl** of saline into the wells used for interference testing

Step 2: Pipet **20 µl** of (diluted) sample into each well

Step 3: Pipet **20 µl** of 1,0 EU/ml Endotoxin Control into the wells

Step 4: Pipet **20 µl** of donor blood into the wells

Continue with Step 5 of the Whole Blood Stimulation-microtiter plate method procedure.

7.2: ELISA Procedure

.

All components must be at room temperature (15-28°C) before use. Do *not* thaw frozen specimens by heating them in a waterbath. The ELISA is carried out at room temperature.

1 For control of the ELISA procedure, the stimulation supernatants of the Endotoxin Controls (EC) and the LTA control are used. P1, P2, etc. are the stimulation supernatants of the test probes.

2 Sample distribution: see Microplate Template below (quantitative assay)

	1	2	3	4	5	6	7	8	9	10	11	12
A	EC 5,0	EC 5.0	EC 5.0	P02	P04	P07	P10	P12	P15	P18	P20	P23
B	EC 2,5	EC 2.5	EC 2.5	P02	P05	P07	P10	P13	P15	P18	P21	P23
C	EC 1.0	EC 1.0	EC 1.0	P02	P05	P08	P10	P13	P16	P18	P21	P24
D	EC 0.5	EC 0.5	EC 0.5	P03	P05	P08	P11	P13	P16	P19	P21	P24
E	EC 0.25	EC 0,25	EC 0.25	P03	P06	P08	P11	P14	P16	P19	P22	P24
F	saline	saline	saline	P03	P06	P09	P11	P14	P17	P19	P22	P25
G	LTA	LTA	LTA	P04	P06	P09	P12	P14	P17	P20	P22	P25
H	P 01	P 01	P01	P04	P07	P09	P12	P15	P17	P20	P23	P25

4 Add **100 µl** Enzyme-Labeled Antibody to every well

3 Pipet **100 µl** of supernatants of Endotoxin Controls, LTA control, those of the

13

negative (saline) control and of the samples into the wells prepared.
Use a disposable-tip micropipet for the samples, changing the tip between samples, to avoid contaminations.
5 Cover the plate and mix for **90 minutes** on a microplate mixer.
6 Decant, then wash. For assays using centrifuged blood supernatants, wash each well 4 times with **300 µl** Buffered Wash Solution. For assays using resuspended blood, wash 5 to 6 times with 300 µl per well.
If this step is performed manually, remove as much moisture as possible during the decanting; this will greatly enhance precision. A technical Data Sheet describing the procedure in detail is available on request.
Before adding the TMB/Substrate solution, tap the plate face down on adsorbant paper to shake off all residual droplets, being careful not to dislodge the strips from the frame.
7 Add **200 µl** of TMB/Substrate Solution to every well.
8 Incubate without shaking for **30 minutes** in the dark.
9 Add **50 µl** of Stop Solution to every well.
Tapping the plate gently after the addition of Stop Solution will aid mixing and improve precision. The Stop Solution is acidic.
Handle carefully, and use safety gloves and eye protection.
10 Read at 450 nm, **within 15 minutes** of adding Stop Solution

8. DATA ANALYSIS AND ASSOCIATED ERRORS

The Endotoxin Controls and a negative control (saline) should routinely be assayed in each run.
The assay should be considered acceptable only if the following criteria are met:

.

The mean OD of the 0.5 EU/ml endotoxin control exhibits an OD that is greater than 1.6x the mean ODf the negative saline control.

The OD of the PPC satisfies the requirement stated in the Interference testing for products.

Interference testing for products

A. Quantitative IPT assay:

There is no interference if

0.5 x the median $EC_{1,0}$ < median interference test < 2x median $EC_{1,0}$
14

If the median assay result (in terms of OD) falls outside the 50 to 200% range of the median of the incubation of the 1,0 EU/ml control in the absence of sample, repeat the experiment using 100 µl of *diluted* sample until it yields an OD reading inside the range. (Dilute the sample with saline, e.g. 1:10, 1:100, etc.). Samples exhibiting interference should be assayed at the lowest dilution not causing interference.

B. Threshold IPT assay:

There is no interference if

0.5 x the median EC $_{0,5}$ median interference test < 2x median EC $_{0,5}$

Interpretation

A. Quantitative IPT assay
???????????

B. Threshold IPT assay (rabbit equivalent test)

The results are given as *positive* or *negative* (non-pyrogenic). A sample is considered positive if the mean OD of the sample is equal or greater than the mean OD of the 0.5 EU/ml standard.

9. PREDICTION MODEL

Rabbits are likely to develop fever if tested with 10ml/kg of the sample if

$$OD_{Sample} > OD_{mean\ (0.5\ EEU)}$$

10. HEALTH SAFETY AND ENVIRONMENT

· For *in vitro* use only.
· Do not use reagents beyond their expiration dates.

Bio-Safety
Human blood has to be considered infectious and handled accordingly.
This kit contains components of human origin which, when tested by FDA-approved
15

methods, were found non-reactive for hepatitis B surface antigen and for HIV antibody. No known tests can guarantee, however, that products derived from human blood will not be infectious. Handle, therefore, as if capable of transmitting infectious agents.
.

Stop Solution and TMB/Substrate Solution
Avoid contact with the Stop Solution, which is acidic. Wear gloves and eye protection. If this reagent comes into contact with skin, wash thoroughly with water and seek medical attention, if necessary. The reagent is corrosive; therefore, the instrument employed to dispense it should be thoroughly cleaned after use. The TMB/Substrate Solution contains peroxide. Since peroxides are strong oxidizing agents, avoid all bodily contact with the TMB/Substrate Solution

16

Institute for Health and Consumer Protection

11. ANNEX (Pipetting scheme for the whole blood assay)

Part 1: Whole blood stimulation (all values in µl)

Tube account	Stimulation sample	saline	Endotoxin Control (0.5 EEU)	Endotoxin Control (1 EEU)	Endotoxin Control (2 EEU)	LTA control	Test sample	Donor blood	Incubate overnight at 37°C	Mix the samples. Centrifuge for 2 minutes at 10000 x g (if necessary). Take 150 µl from the supernatant. Test immediately with the ELISA system or store at-20 °C.
3	Endotoxin Control (0.5 EEU)	1000	100	-	-		-	100		
2	Endotoxin Control (1 EEU)	1000	-	100	-		-	100		
2	Endotoxin Control (2 EEU)	1000	-	-	100		-	100		
2	LTA control	1000				100		100		
3	Blank (0)	1100						100		
3	Interference test,	900	100	-	-		100 (diluted)	100		
3	Test samples 1, 2, 3	1000	-	-	-		100	100		

3

Part 2: Procedure (all values in µl)

Well	Supernatants from Stimulation	Enzyme-labeled Antibody		Substrate		Stop solution	
D2, E2, F2 (Blank)	100	150	Incubate 90 min at RT on a plate mixer at 350-400 rpm. Decant.Wash 4 times with 300 µl Buffered Wash Solution	200	Incubate 30 min at RT	50	Read at 450 nm
EC: see template schedule	100	150		200		50	
G1, G2 (LTA control)	100	150		200		50	
Samples: see template schedule	100	150		200		50	

4

12. REFERENCES

1 Hartung T., Wendel A. : Detection of pyrogens using human whole blood. In Vitro Toxicology; 9(4): 353-59.
2 Fennrich S., Fischer M., Hartung T., Lexa P., Montag-Lessing T., Sonntag H.-G., Weigand M. und Wendel A. : Detection of endotoxins and other pyrogens using human whole blood . Dev Biol Stand.Basel , 1999, 101:131-39.
3 Hartung T., Aaberge I., Berthold S., Carlin G., Charton E., Coecke S., Fennrich S., Fischer M., Gommer M., Halder M., Haslov K., Montag-Lessing T., Poole S., Schechtman L., Wendel A. und Werner-Felmayer, G. : ECVAM workshop on novel pyrogen tests based on the human fever reaction. ATLA 2001, 29:99-123.
4 Morath S., Geyer A., Hartung T. : Structure-function relationship of cytokine induction by lipoteichoic acid from Staphylococcus aureus. J. Exp. Med., 2001, 193:393-397.
5 Bonenberger J., Diekmann W., Fennrich S., Fischer M., Friedrich A., Hansper M., Hartung T., Jahnke M., Löwer J., Montag-Lessing T., Petri E., Sonntag H.-G., Weigand M., Wendel A. und Zucker B. : Pyrogentestung mit Vollblut-Zusammenfassung eines Status-Workshops am Paul-Ehrlich-Institut, Langen, am 22.11.99. Bundesgesundheitsbl-Gesundheitsforsch-Gesundheitsschutz 2000, 43:525-533.
6 Petri E., van de Ploeg A., Habermaier B. and Fennrich S.: Improved detection of pyrogenic substances on polymer surfaces with an ex vivo human whole-blood assay in comparison to the Limulus amoebocyte lysate test. In: Balls M., van Zeller A.-M. and Halder M.: Progress in the reduction, refinement and replacement of animal experimentation. Elsevier 2000, 339-345.
7 Fennrich S., Wendel A. and Hartung T.: New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 1999, 16:146-149.

[This Page Intentionally Left Blank]

Validation of Biomedical Testing Methods

Human Whole Blood Pyrogen Test

Standard Operating Procedure for the Validation Phase

Only the responsible of the GLP/QA Unit is allowed to make copies of this document. Extra examples can be obtained at the GLP/QA Unit. Quality Documents are only valid if they are signed by the responsible of the GLP/QA Unit in and provided with a blue copy number	Copy number

Sop-WBT-KNv02

JOINT RESEARCH CENTRE
EUROPEAN COMMISSION
Institute for Health and Consumer Protection
ECVAM

Human Whole Blood Pyrogen Test

Standard Operating Procedure

Drafted by:	*Name* *Date* *Signature*	*Thomas Hartung,* *30/09/02*	*M. D. Ph.D.*
Reviewed by:	*Name* *Date* *Signature*		
Approved by:	*Name* *Date* *Signature*		
Issued by:	*Name* *Date* *Signature*		

*Owner/Trainer: Signature: Date:

PAGE OF CHANGES

Date of change/ Date of draft:	Version-number:	Changed page(s):	Summary of the change(s):	Changed by/Sign.:
01.04.01				
29.06.01		10		
20.09.01		2, 4, 5, 6, 8, 10, 11, 12, 13, 14 16		
14.12.01		2, 6, 8, 9, 10, 11		
23.02.02		2, 5, 6,7, 8, 9 10 11		
28.2.02		2, 6, 7, 10, 12		
06.03.02		2, 6, 7, 8, 9, 10, 11,		
30.09.02		12, 14, 17, 1, 4, 6, 7, 9, 10, 11, 12, 13, 14,		

TABLE OF CONTENTS Page no.

2. PURPOSE 6
3. SCOPE / LIMITATIONS 7
4. METHOD OUTLINE 8
5. DEFINITIONS / ABBREVIATIONS 9
6. MATERIALS 10
6.1. Materials required and not provided 10
6.1.1 Materials for Blood Incubation 10
6.1.2 Materials for ELISA procedure 11
6.2. Materials Supplied in ELISA kit 11
7. METHODS 14
7.1. Blood Incubation 14
7.2. ELISA Procedure 19
MINIMUM ASSAY SUITABILITY REQUIREMENTS 21
8. HEALTH SAFETY AND ENVIRONMENT 21
9. ANNEX (Pipetting scheme for the whole blood assay) 23
10. REFERENCES 25

THIS SOP WAS AMENDED FOR THE VALIDATION PHASE ONLY. IT DOES THEREFORE ONLY REPLACE THE PREVIOUS VERSION FOR THIS SERIES OF EXPERIMENTS.

1. INTRODUCTION

The whole blood pyrogen test (in vitro pyrogen test IPT) is a two-part assay for the detection of pyrogenic contamination. It involves incubation of the sample with human blood, followed by an enzyme immunoassay for the measurement of IL-1b.

A pyrogen is a substance that causes fever. Bacterial contaminations, which contain exogenous pyrogens, can be deadly. This problem is of great significance for drug safety.

Also, medical devices and biologically produced substances obtained from bacteria and other microorganisms may cause release of endogenous pyrogens (e.g., IL-1b).

Exogenous pyrogens include metabolic substances and cell-wall components of microorganisms. These substances are present during the "normal" course of an infectious disease. Infections by gram-negative and gram-positive bacteria are equal in frequency. Both of these bacterial types can activate the release of endogenous pyrogens, which cause fever through the thermoregulatory center in the brain. Although these reactions can occur during the "normal" course of an infectious disease, a deadly shock syndrome can occur in the worst case.

Due to these risks, product safety legislation demands rigorous quality checks for pyrogenic contamination of drugs and devices intended for parenteral use. For example, testing in rabbits for medical end products is required in Germany. Products in development and a few end products are allowed to be controlled by the Limulus assay. The first pyrogen assay, based on human whole blood stimulation by pyrogens, was developed by Hartung et al. (3,4).

2. PURPOSE

This assay simulates *in vitro* the normal human reaction to exogenous pyrogens. A few drops of human blood are mixed with the sample, and exogenous pyrogens in the sample are recognized by immunocompetent cells in the human blood.
These cells release IL-1b, which is measured by an integrated ELISA system.

3. SCOPE / LIMITATIONS

Limit of detection is ≤ 0,25 EEU/ml, not suitable for test samples interfering with blood cytokine release.

THIS SOP WAS AMENDED FOR THE VALIDATION PHASE ONLY. IT DOES THEREFORE ONLY REPLACE THE PREVIOUS VERSION FOR THIS SERIES OF EXPERIMENTS.

4. METHOD OUTLINE

The procedure has two parts:

1. Incubation of the sample with (diluted) human blood
2. An enzyme immunoassay for the measurement of IL-1b.

Ad 1) Blood incubation

Diluted human whole blood is incubated for 10-24 hours together with saline and the sample in pyrogen-free reaction tubes and the supernatant is taken off for further examination.

Ad 2) Capture of Endogenous Pyrogens (ELISA procedure)

Samples (supernatants of blood stimulation) are distributed into the wells of a microplate which are coated with monoclonal antibodies specific for IL-1b.

An enzyme-conjugated polyclonal antibody against IL-1b is added. During a 90-minute incubation, a sandwich complex consisting of two antibodies and the IL-1b is formed. Unbound material is removed by a wash step.

A chromogenic substrate (3,3´,5,5´ -tetramethylbenzidine, TMB) reactive with the enzyme label is added. Color development is terminated by adding a stop solution after 30 minutes. The resulting color, read at 450 nm, is directly related to the IL-1b concentration. Bi-chromatic measurement with a 600-690 nm reference filter is recommended.

5. DEFINITIONS / ABBREVIATIONS

The following abbreviations are used in this work-book.

Ab	antibody
°C	degrees Celsius (Centigrade)
EC	endotoxin control
EEU	endotoxin equivalent unit
ELISA	Enzyme-Linked ImmunoSorbent Assay
EU	endotoxin unit of the international standard
h	hour
H_2SO_4	sulphuric acid
IL	interleukin
LPS	lipopolysaccharide (exogenous pyrogen from Gram-negative bacteria)
LTA	lipoteichoic acid (exogenous pyrogen from Gram-positive bacteria)
µl	microlitre
mg	milligram
ml	millilitre
min	minute
MVD	maximum valid dilution
NaCl	sodium chloride, 0,9%
nm	nanometre
NPC	negative product control
PPC	positive product control
OD	optical density
rpm	rounds per minute
RT	room temperature
TMB	3,3´,5,5´-Tetramethylbenzidine
WDB	wash/dilution buffer
x g	x gravity

6. MATERIALS

6.1. Materials required and not provided

The components listed below are recommended, but equivalent devices may also be used: it is the users responsibility to validate the equivalence.
For all steps excluding the ELISA procedure sterile and pyrogen-free materials have to be used (e.g. tips, containers, solutions).

6.1.1 Materials for Blood Incubation

Equipment

· Incubator or thermoblock (37°C ± 1°C)
· Multipette or adjustable 100 to 1000 µl pipetters
· Centrifuge (recommended)
· Vortex mixer

Consumables

· Heparinized tubes for blood sampling (e.g. Sarstedt S-MONOVETTE 7.5 ml, 15 IU/ml Li-Heparin)
· Sarstedt multifly needle set, pyrogenfree, for S-Monovette
· 1.5 ml closable, pyrogen-free reaction tubes
· Reservoir for saline
· 12 ml (PS)or 15 ml (PP) tubes from greiner bio-one for dilution of substances
· Sterile and pyrogen-free tips 100 µl and 1000 µl **or**
· Combitips for multipette, 10 ml and 2,5 ml for pipetting saline and blood

6.1.2 Materials for ELISA procedure

Equipment

· Multichannel pipettor

· Microplate mixer

· Microplate washer

· Microplate reader capable of readings at 450 nm (optional reference filter in the range of 600-690 nm)

· A software package for facilitating data generation, analysis, reporting, and quality control

Consumables

· Graduated cylinder and plastic storage container for Buffered Wash Solution

· Tip-Tubs for reagent aspiration with Multichannel pipettor

6.2. Materials Supplied in ELISA kit

Components supplied in that kit are *not* interchangeable with other lots of the same components.

IL-1b Ab-coated Microplate: One 96-well polystyrene microplate, packaged in a zip-lock foil bag, with desiccant. The plate consists of twelve strips mounted in a frame. Each strip includes eight anti-IL-1b Ab-coated wells. Additionally, individual wells can be separated from the strip to enable the complete use of all the wells of a kit. Well positions are indexed by a system of letters and numbers (A through H, 1

through 12) embossed on the left and top edges of the frame. Store refrigerated: stable at 2-8°C until the expiration date marked on the label.

Enzyme-Labeled Antibody : One amber vial containing 16 ml of liquid reagent, ready-to-use. The reagent contains horseradish peroxidase-labeled, affinity-purified, polyclonal (rabbit) anti-IL-1b antibodies, with preservative. Store refrigerated: stable at 2-8°C for 30 days after opening, or until the expiration date marked on the label. *Do not freeze.*

Endotoxin Control: One vial of an endotoxin control. The control is supplied lyophilized. Store refrigerated: stable at 2-8°C until the expiration date marked on the label. At least 30 minutes before use, reconstitute control vial with saline. Prepare serial dilutions in saline (see 7. Methods). Mix by vortexing. After preparation, the stock solution can be stored (see 7. Methods).

Saline: Three glass vials, each containing pyrogen-free saline. This is intended for the dilution of donor blood samples and for reconstitution of the Endotoxin Control. Store refrigerated: stable at 2-8°C until the expiration date marked on the label. Use immediately after opening and discard unused volumes.

TMB/Substrate Solution: Two amber vials, each containing 11 ml of a buffered reagent, ready-to-use. The reagent contains a hydrogen peroxide substrate and 3,3´,5,5´-tetramethylbenzidine (TMB). Store refrigerated and protected from light: stable at 2-8°C until the expiration date marked on the label. *Do not freeze*.

Buffered Wash Solution Concentrate: One vial containing 75 ml of a concentrated (10X) buffered saline solution, with surfactants and preservative. Using a transfer container, dilute the contents of the vial with **675 ml** distilled or deionized water for a total volume of **750 ml**.

Store refrigerated: stable at 2-8°C for 7 days after preparation, or until the expiration date marked on the label.

Stop Solution: One vial containing an acidic solution, for terminating the color reaction. The reagent is supplied ready-to-use. Handle with care, using safety gloves

and eye protection. Store refrigerated: stable at 2-8°C for 8 weeks after opening, or until the expiration date marked on the label.

7. METHODS

7.1. Blood Incubation

Blood Collection

Collect blood by venipuncture into heparinized tubes. The blood collection system must be pyrogen-free. The procedure calls for 100 µl of heparinized whole blood per reaction tube. The blood can be stored in the collection tube at room temperature (15-28°C) for 4 hours. Incubation of the sample should be started within this time.

Note:

1. Blood donors should show no evidence of disease or need of medication during the last two weeks.
2. Each assay should include the Endotoxin Controls and the saline control in quadruplicate.
3. Use disposable tip pipets to avoid contamination of reagents and samples.
4. During ELISA procedure, the wells should be washed carefully.
5. The test samples should be done in quadruplicate.
6. The contents of the wells must be decanted or aspirated completely before pipetting wash solution.
7. Deviations from the procedure (incubation time/temperature) may cause erroneous results. The ELISA procedure should be run without interruption. Diluted samples should be tested within an hour.

Storage of the substances

- please keep all substances at 4°C

Spiking of the substances

Part 1)

5 blinded spikes have been sent out by PEI

They are bearing a code for

a) the respective drug
b) the test method, in this case WBT-KN
c) a random blinding number

- please pipet 500 µl of the respective substance into an Eppendorf tube
- vortex the respective vial with the blinded spike for about 5 seconds
- add 25 µl of the spike to the substance and vortex for another 5 seconds
- perform the dilutions according to the instructions below

Dilution of the substances

- for dilution, please use either 12 ml or 15 ml tubes from greiner bio-one
- each substance has to be vortexed for about 5 seconds immediately before performing Step 3 of the Whole Blood Stimulation.

Substance 1: Glucose 5%

Maximum valid dilution =1:75; add 40 µl of substance to 2960 µl of saline

Substance 2: EtOH 13%

Maximum valid dilution = 1:37.5 ; add 80 µl of substance to 2920 µl of saline

Substance 3: MCP

Maximum valid dilution: 1:375; add 8 µl of substance to 2992 µl of saline

Substance 4: Orasthin

Maximum valid dilution: 1:750: add 4 µl of substance to 2996 µl of saline

Substance 5: Binotal

Maximum valid dilution: 1:150; add 20 µl of substance to 2980 µl of saline

Substance 6: Fenistil

Maximum valid dilution: 1:187.5; add 16 µl of substance to 2984 µl of saline

Substance 7: Sostril

Maximum valid dilution: 1:150; add 20 µl of substance to 2980 µl of saline

Substance 8: Beloc

Maximum valid dilution: 1:150; add 20 µl of substance to 2980 µl of saline

Substance 9: Drug A

Maximum valid dilution: 1:37.5; add 80 µl of substance to 2920 µl of saline

Substance 10: Drug B

Maximum valid dilution: 1:75; add 40 µl of subsubstance to 2960 µl of saline

Part 2)

(unblinded)

- Positive Product Control (PPC)

dilute the respective substance according to the instructions above

vortex for about 5 seconds

pipet 500 µl of the diluted substance into an Eppendorf tube

add 25 µl of the unblinded PPC-LPS spike handed out by PEI

- Negative Product Control (NPC)
dilute the respective substance according to the instructions above
vortex for about 5 seconds
pipet 500 µl of the diluted substance into an Eppendorf tube
add 25 µl of saline

Endotoxin dilution for the Dose-Response Curve

IPT assays must include the 0.5 EU/ml + saline control in quadruplicate.

Dissolve the contents of the vial containing O113 provided by NIBSC with 5 ml of saline yielding a stock solution of 2000 EU/ml.

EC = Endotoxin Control, for use in the assay.

Solution	amount added to saline	Volume of saline	Resulting solution
Stock (2000 EU/ml)	100 µl	900 µl	200 EU/ml
200 EU/ml	100 µl	900µl	20 EU/ml
20 EU/ml	100 µl	900 µl	2 EU/ml
2 EU/ml	500 µl	500 µl	1 EU/ml (EC)
1 EU/ml	500 µl	500 µl	0,5 EU/ml (EC)

The stock solution of the Endotoxin Standard may be aliquoted (e.g. 100 µl aliquots) and kept at –20 °C for up to 6 months.

Whole Blood Stimulation

Perform incubation of blood samples in 1.5 ml pyrogen-free reaction tubes. Preferably, use a laminar-flow bench**.** All consumables and solutions have to be sterile and pyrogen-free.

Step 1: Draw up an incubation plan according to the template below

Step 2: Add **1000 µl** saline into each reaction tube.

Step 3: Add **100 µl** of Endotoxin Controls and negative saline control or samples in quadruplicate into the respective reaction tubes according to the prepared incubation plan.

Step 4: Add **100 µl** of donor blood, mixed by gentle inversion, into each reaction tube.

Step 5: Close the tubes and invert them once or twice before starting the incubation.

Step 6: Incubate the closed reaction tubes in an incubator or a heating block overnight (10-24 hours) at 37°C ± 1°C.

Step 7: Mix the incubation tubes thoroughly by inverting the tubes. Incubations are to be centrifuged for 2 minutes at 10.000 g and the clear supernatant is used for the ELISA procedure. Take aliquots of ≥ **150 µl.**

The supernatants can be tested immediately by the ELISA System or may be stored at -20°C for testing at a later time.

Freeze additional aliquots.

7.2. ELISA Procedure

All components must be at room temperature (15-28°C) before use. Do *not* thaw frozen specimens by heating them in a waterbath. The ELISA is carried out at room temperature.

1 For control of the ELISA procedure, the stimulation supernatants of the Endotoxin Controls (EC) are used.

2 Sample distribution: see Microplate Template below.

A	NPC (A)	NPC (A)	PPC (A)	PPC (A)	PPC (A)	PPC (A)	1 (B)	1 (B)	1 (B)	1 (B)	2 (C)	2 (C)
B	NPC (A)	NPC (A)	1 (A)	1 (A)	1 (A)	1 (A)	2 (B)	2 (B)	2 (B)	2 (B)	2 (C)	2 (C)
C	EC 1,0	EC 1,0	2 (A)	2 (A)	2 (A)	2 (A)	3 (B)	3 (B)	3 (B)	3 (B)	3 (C)	3 (C)
D	EC 1,0	EC 1,0	3 (A)	3 (A)	3 (A)	3 (A)	4 (B)	4 (B)	4 (B)	4 (B)	3 (C)	3 (C)
E	EC 0,5	EC 0,5	4 (A)	4 (A)	4 (A)	4 (A)	5 (B)	5 (B)	5 (B)	5 (B)	4 (C)	4 (C)
F	EC 0,5	EC 0,5	5 (A)	5 (A)	5 (A)	5 (A)	NPC (C)	NPC (C)	NPC (C)	NPC (C)	4 (C)	4 (C)
G	saline	saline	NPC (B)	NPC (B)	NPC (B)	NPC (B)	PPC (C)	PPC (C)	PPC (C)	PPC (C)	5 (C)	5 (C)
H	saline	saline	PPC	PPC	PPC	PPC	1	1	1	1	5	5

			(B)	(B)	(B)	(B)	(C)	(C)	(C)	(C)	(C)	(C)

A, B, C : e.g. Substances 1, 2, 3

1-5 : blinded spikes 1-5

EC : Endotoxin Control

NPC: negative product control

PPC: positive product control

3 Add **100 µl** Enzyme-Labeled Antibody to every well

4 Pipet **100 µl** of supernatants of Endotoxin Controls, those of the negative saline control and of the samples into the wells prepared.

Use a disposable-tip micropipet for the samples, changing the tip between each sample and control, to avoid contaminations.

5 Mix for **90 minutes** on a microplate mixer at 350-400 rpm.

6 Decant, then wash. For assays using centrifuged blood supernatants, wash each well 4 times with **300 µl** Buffered Wash Solution. For assays using resuspended blood, wash 5 to 6 times with 300 µl per well.

If this step is performed manually, remove as much moisture as possible during the decanting by inverting the washed microplate and tapping out the residual washing buffer on blotting paper or a paper towel, being careful not to dislodge the strips from the frame.

7 Add **200 µl** of TMB/Substrate Solution to every well.

8 Incubate without shaking for **30 minutes** in the dark.

9 Add **50 µl** of Stop Solution to every well.

*Tapping the plate **gently** after the addition of Stop Solution will aid mixing and improve precision. The Stop Solution is acidic.*

Handle carefully, and use safety gloves and eye protection.

10 Read at 450 nm, **within 15 minutes** of adding Stop Solution. Bi-chromatic measurement with a reference wavelength of 600-690 nm is recommended.

MINIMUM ASSAY SUITABILITY REQUIREMENTS

The assay should be considered acceptable only if the following minimum criteria are met:

The mean optical density of the 0.5 EU/ml endotoxin control exhibits an OD that is greater than 1.6 times the mean optical density of the negative saline control.

8. HEALTH SAFETY AND ENVIRONMENT

· For *in vitro* use only.

· Do not use reagents beyond their expiration dates.

Bio-Safety

Human blood has to be considered infectious and handled accordingly.

Stop Solution and TMB/Substrate Solution

Avoid contact with the Stop Solution, which is acidic. Wear gloves and eye protection. If this reagent comes into contact with skin, wash thoroughly with water and seek medical attention, if necessary. The reagent is corrosive; therefore, the instrument employed to dispense it should be thoroughly cleaned after use. The TMB/Substrate Solution contains peroxide. Since peroxides are strong oxidizing agents, avoid all bodily contact with the TMB/Substrate Solution

9. ANNEX (Pipetting scheme for the whole blood assay)

Part 1: Whole blood stimulation (all values in µl)

Tube account	Stimulation sample	saline	Endotoxin Control (0.5 – 1.0 EU/ml)	Test sample	Donor blood		
4	Endotoxin Control (0.5 – 1.0 EU/ml)	1000	100	-	100	Incubate overnight at 37°C	Mix the samples. Centrifuge for 2 minutes at 10000 x g (if necessary). Take 150 µl from the supernatant. Test immediately with the ELISA system or store at -20 °C.
4	Blank (0)	1100			100		
4	Test samples (1-8)	1000	-	100	100		

Part 2: ELISA procedure (all values in µl)

Well	Supernatants from Stimulation	Enzyme-labeled Antibody		Substrate		Stop solution	
G/H 1/2 (Blank)	100	100	Incubate 90 min at RT on a plate mixer at 350-400 rpm. Decant.Wash 4 times with 300 µl Buffered Wash Solution	200	Incubate 30 min at RT	50	Read at 450 nm (600-690 nm reference wave-length recom-mended)
EC: B-F 1/2	100	100		200		50	
Samples: see template schedule	100	100		200		50	

10. REFERENCES

1. Hartung T., Wendel A. : Detection of pyrogens using human whole blood. In Vitro Toxicology; 9(4): 353-59.
2. Fennrich S., Fischer M., Hartung T., Lexa P., Montag-Lessing T., Sonntag H.-G., Weigand M. und Wendel A. : Detection of endotoxins and other pyrogens using human whole blood . Dev Biol Stand.Basel , 1999, 101:131-39.
3. Hartung T., Aaberge I., Berthold S., Carlin G., Charton E., Coecke S., Fennrich S., Fischer M., Gommer M., Halder M., Haslov K., Montag-Lessing T., Poole S., Schechtman L., Wendel A. und Werner-Felmayer, G. : ECVAM workshop on novel pyrogen tests based on the human fever reaction. ATLA 2001, 29:99-123.
4. Morath S., Geyer A., Hartung T. : Structure-function relationship of cytokine induction by lipoteichoic acid from Staphylococcus aureus. J. Exp. Med., 2001, 193:393-397.
5. Bonenberger J., Diekmann W., Fennrich S., Fischer M., Friedrich A., Hansper M., Hartung T., Jahnke M., Löwer J., Montag-Lessing T., Petri E., Sonntag H.-G., Weigand M., Wendel A. und Zucker B. : Pyrogentestung mit Vollblut-Zusammenfassung eines Status-Workshops am Paul-Ehrlich-Institut, Langen, am 22.11.99. Bundesgesundheitsbl-Gesundheitsforsch-Gesundheitsschutz 2000, 43:525-533.
6. Petri E., van de Ploeg A., Habermaier B. and Fennrich S.: Improved detection of pyrogenic substances on polymer surfaces with an ex vivo human whole-blood assay in comparison to the Limulus amoebocyte lysate test. In: Balls M., van Zeller A.-M. and Halder M.: Progress in the reduction, refinement and replacement of animal experimentation. Elsevier 2000, 339-345.
7. Fennrich S., Wendel A. and Hartung T.: New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 1999, 16:146-149.

[This Page Intentionally Left Blank]

Appendix A2

The Human WB/IL-1β *In Vitro* Pyrogen Test: Application of Cryopreserved (Cryo) Human WB

ECVAM Background Review Document (March 2006) A-135
ECVAM Standard Operating Procedure for the Catch-Up Validation Phase (July 2004)..... A-215

[This Page Intentionally Left Blank]

BRD: CRYO WB/IL-1 March, 2006

THE HUMAN WHOLE BLOOD/IL-1 IN VITRO PYROGEN TEST (CRYO WB/IL-1)

Use of cryopreserved human whole blood

BRD: CRYO WB/IL-1 March, 2006

Contents

1 RATIONALE FOR THE PROPOSED TEST METHOD 4
1.1 INTRODUCTION 4
1.2 REGULATORY RATIONALE AND APPLICABILITY 5
1.3 SCIENTIFIC BASIS FOR THE PROPOSED TEST METHOD 6
2 TEST METHOD PROTOCOL COMPONENTS 8
2.1 OVERVIEW OF TEST METHOD. 8
2.2 RATIONAL FOR SELECTED TEST COMPONENTS 9
2.3 BASIS FOR SELECTION OF THIS TEST METHOD 13
2.4 PROPRIETARY COMPONENTS 13
2.5 REPLICATES 14
2.6 MODIFICATIONS APPLIED AFTER VALIDATION 14
2.7 DIFFERENCES WITH SIMILAR TEST METHODS 14
3 SUBSTANCES USED FOR VALIDATION 15
3.1 SELECTION OF SUBSTANCES USED 15
3.2 NUMBER OF SUBSTANCES 15
3.3 DESCRIPTION OF SUBSTANCES USED 16
3.4 SAMPLE CODING PROCEDURE 17
3.5 RECOMMENDED REFERENCE CHEMICALS 17
4 *IN VIVO* REFERENCE DATA ON ACCURACY 18
4.1 TEST PROTOCOL *IN VIVO* REFERENCE TEST METHOD. 18
4.2 ACCURACY 18
4.3 ORIGINAL RECORDS 19
4.4 QUALITY OF DATA 19
4.5 TOXICOLOGY 20
4.6 BACKGROUND ON ASSAY PERFORMANCE 20
5 TEST METHOD DATA AND RESULTS 21
5.1 TEST METHOD PROTOCOL 21
5.2 ACCURACY AND RELIABILITY 22
5.3 STATISTICS 24
5.4 TABULATED RESULTS 28
5.5 CODING OF DATA 30
5.6 CIRCUMSTANCES 30
5.7 OTHER DATA AVAILABLE 30
6 TEST METHOD ACCURACY 31
6.1 ACCURACY 31
6.2 CONCORDANCY TO *IN VIVO* REFERENCE METHOD 34
6.3 COMPARISON WITH REFERENCE METHODS 34
6.4 STRENGTH AND LIMITATIONS 34
6.5 DATA INTERPRETATION 34
6.6 COMPARISON TO OTHER METHODS 35
7 TEST METHOD RELIABILITY (REPEATABILITY/REPRODUCIBILITY) 36
7.1 SELECTION OF SUBSTANCES 36
7.2 RESULTS 36
7.3 HISTORICAL DATA 38
7.4 COMPARISON TO OTHER METHODS 38
8 TEST METHOD DATA QUALITY 39
8.1 CONFORMITY 39
8.2 AUDITS 39
8.3 DEVIATIONS 39
8.4 RAW DATA 39
9 OTHER SCIENTIFIC REPORTS AND REVIEWS 40
9.1 SUMMARY 40
9.2 DISCUSSION 41

9.3 RESULTS OF SIMILAR VALIDATED METHOD 42
10 ANIMAL WELFARE CONSIDERATIONS (REFINEMENT, REDUCTION, AND REPLACEMENT) 45
10.1 DIMINISH ANIMAL USE 45
10.2 CONTINUATION OF ANIMAL USE 45
11 PRACTICAL CONSIDERATIONS 47
11.1 TRANSFERABILITY 47
11.2 TRAINING 47
11.3 COST CONSIDERATIONS 48
11.4 TIME CONSIDERATIONS 48
12 REFERENCES 49
13 IN VITRO PYROGEN TEST WITH CRYOPRESERVED HUMAN WHOLE BLOOD ACCORDING THE KONSTANZ METHOD (KN CRYO WB/IL-1). 51
13.1 RATIONALE 51
13.2 TEST METHOD PROTOCOL COMPONENTS 51
13.3 SUBSTANCES USED FOR VALIDATION 51
13.4 PRELIMINARY ESTIMATE OF THE TEST METHOD ACCURACY 51
13.5 TEST METHOD ACCURACY 55
13.6 TEST METHOD RELIABILITY (REPRODUCIBILITY) 59
13.7 SUMMARY AND CONCLUSION 59
14 SUPPORTING MATERIALS (APPENDICES) 60
14.1 STANDARD OPERATING PROCEDURE (SOP) OF THE PROPOSED METHOD 60
14.2 STANDARD OPERATING PROCEDURE (SOP) OF THE REFERENCE METHOD 60
14.3 PUBLICATIONS 60
14.4 ORIGINAL DATA 66
14.5 PERFORMANCE STANDARDS 66

Appendix A – Method protocol(s) and trial plan(s)

Appendix B – Hardcopies of relevant publications

Appendix C – List of abbreviations and definitions

Appendix D –Intralaboratory reproducibility of CRYO WB/IL-1, KN CRYO WB/IL-1, 96-wells WB/IL-1

1 Rationale for the Proposed Test Method

1.1 Introduction

1.1.1. *Describe the historical background for the proposed test method, from original concept to present. This should include the rationale for its development, an overview of prior development and validation activities, and, if applicable, the extent to which the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards.*

Pyrogens, a chemically heterogeneous group of hyperthermia- or fever-inducing compounds, derive from bacteria, viruses, fungi. Subjects react to such microbial products during an immune response by producing endogenous pyrogens such as prostaglandins and the pro-inflammatory cytokines interleukin-1 (IL-1), interleukin-6 (IL-6) and tumour necrosis factor-α (TNF-α) (Dinarello, 1999). Depending on the type and amount of pyrogen challenge and the sensitivity of an individual, even life-threatening shock-like conditions can be provoked. To assure quality and safety of any pharmaceutical product for parenteral application in humans, pyrogen testing is therefore imperative.

Depending on the medicinal product, one of two animal-based pyrogen tests is currently prescribed by the respective Pharmacopoeias, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET). For the rabbit pyrogen test, sterile test substances are injected intravenously to rabbits and any rise in body temperature is assessed. This *in vivo* test detects various pyrogens but not alone the fact that large numbers of animals are required to identify a few batches of pyrogen-containing samples argues against its use. In the past two decades, the declared intention to refine, reduce and replace animal testing, has lowered rabbit pyrogen testing by 80% by allowing to use the BET as an *in vitro* alternative pyrogen test for certain medicinal products (Cooper et al, 1971).

Bacterial endotoxin, comprising largely lipopolysaccharide (LPS) from the cell wall of Gram-negative bacteria that stimulates monocytes/macrophages via interaction with CD14 and toll-like receptor 4 (TLR4) (Beutler and Rietschel, 2003), is the pyrogen of major concern to the pharmaceutical industry due to its ubiquitous sources, its stability and its high pyrogenicity. With the BET, endotoxin is detected by its capacity to coagulate the amoebocyte lysate from the haemolymph of the American horseshoe crab, *Limulus polyphemus*, a principle recognised some 40 years ago (Levin and Bang, 1964). In the US, *Limulus* crabs are generally released into nature after drawing about 20% of their blood and therefore most of these animals survive. However, the procedure still causes mortality of about 30,000 horseshoe crabs per year, which adds to even more efficient threats of the horseshoe crab population like its use as bait for fisheries, habitat loss and pollution (http://www.horseshoecrab.org/).

As with the rabbit test the general problem of translation of the test results to the human fever reaction persists. Moreover, although being highly sensitive, the failure of the BET to detect non-endotoxin pyrogens as well as its susceptibility to interference by e.g. high protein or lipid levels of test substances or by glucans impedes full replacement of the rabbit pyrogen test. Hence, hundreds-of-thousands rabbits per year are still used for pyrogen testing.

A test system that combines the high sensitivity and *in vitro* performance of the BET test with the wide range of pyrogens detectable by the rabbit pyrogen test is therefore required in order to close the current testing gap for pyrogen and to avoid animal-based tests. With this intention and due to improved understanding of the human fever reaction (Dinarello, 1999), test systems based on *in vitro* activation of human monocytes were developed. First efforts date back about 20 years, when peripheral blood mononuclear cells (PBMC) were used to detect endotoxin by monitoring the release of pyrogenic cytokines (Duff & Atkins, 1982; Dinarello et al, 1984). Meanwhile, a number of different test systems, using either whole blood, peripheral blood mononuclear cells (PBMCs) or the monocytoid cell lines MONO MAC 6 (MM6) or THP-1 as a source for human monocytes and various read-outs were established (Poole et al., 1988; Ziegler et al, 1988; Tsuchiya et al, 1980; Hartung & Wendel, 1996; Hartung et al, 2001; Poole et al, 2003). These test systems were validated with the aim of developing a tool for formal inclusion into Pharmacopoeias, an important basis for implementing novel alternative pyrogen tests for product-specific validation.

1.1.2 *Summarize and provide the results of any peer review conducted to date and summarize any ongoing or planned reviews.*
All of the five methods are currently under peer review of the ECVAM Scientific Advisory Committee.

1.1.3 *Clearly indicate any confidential information associated with the test method; however, the inclusion of confidential information is discouraged.*
This document does not contain any confidential information.

1.2 Regulatory rationale and applicability

1.2.1 *Describe the current regulatory testing requirement(s) for which the proposed test method is applicable.*
To assure quality and safety of pharmaceutical products for parenteral application in humans, pyrogen testing is imperative. Depending on the drug, one of two pyrogen tests is currently prescribed by the European Pharmacopoeia, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET), and other national and international guidelines.

1.2.2 *Describe the intended regulatory use(s) (e.g., screen, substitute, replacement, or adjunct) of the proposed test method and how it will be used to substitute, replace, or complement any existing regulatory testing requirement(s).*
Dependent on the product and the presence of relevant clinical data on unexpected pyrogenicity of clinical lots, the proposed test method may be an alternative method for pyrogen testing, thus substituting the rabbit pyrogen test or the BET. In certain cases, the proposed test method may function as a supplementary test method to assess compliance to the licensing dossier.
In case the proposed test method is an alternative for pyrogenicity testing, a thorough cross-validation between the proposed test method and the original method for the specific medicinal product is warranted. In case the proposed test method is an adjunctive test to screen for (unexpected) pyrogenic lots, alert and alarm limits may be established based on consistency of production lots or (preferably) based on actual clinical data.

1.2.3 *Where applicable, discuss the similarities and differences in the endpoint measured in the proposed test method and the currently used in vivo reference test method and, if appropriate, between the proposed test method and a comparable validated test method with established performance standards.*
The current *in vivo* method (rabbit test), as described in the pharmacopoeia, and the proposed *in vitro* test method each determine very different end-points, though the biochemical origins of the response are similar.
The *in vivo* method more resembles a black box, and determines the total rise in body temperature (fever induction) of the animals subjected to the medicinal product, as a result of pyrogens (if any) present in the product.
The proposed test method CRYO WB/IL-1 is an *in vitro* model for the fever response mechanism. It determines the release of cytokines by monocytoid cells into the culture medium upon the interaction of pyrogens and specific receptors on the monocytoid cells. It is these cytokines that trigger the fever response *in vivo*.
Main differences between the *in vivo* and *in vitro* methods are that the latter is quantitative and uses cells of human origin, thus better reflecting the physiological situation.

1.2.4 *Describe how the proposed test method fits into the overall strategy of hazard or safety assessment. If a component of a tiered assessment process, indicate the weight that should be applied relative to other measures.*
The proposed test method CRYO WB/IL-1 may be applied for those medicinal products for which the rabbit test is the only or most reliable method for pyrogenicity testing, since a) the medicinal product is not compatible with the BET or b) the medicinal product contains pyrogens other than Gram-negative endotoxin.
Limit concentrations for pyrogens are established based on consistency lots or actual clinical data or, in the case of endotoxin the endotoxin limit concentration (ELC) as defined for many medicinal products.

1.3 Scientific basis for the proposed test method

1.3.1 *Describe the purpose and mechanistic basis of the proposed test method.*
The proposed *in vitro* method is intended to determine the presence of pyrogens in medicinal products for parenteral use. The proposed test method is an *in vitro* model of the human fever response. It determines the release of cytokines upon the interaction of pyrogens and specific Toll-like receptors on the monocytoid cells (Beutler and Rietschel, 2003). These cytokines trigger the fever response *in vivo*.

1.3.2 *Describe what is known and not known about the similarities and differences of modes and mechanisms of action in the proposed test method as compared to the species of interest (e.g., humans for human health-related toxicity testing).*
An important feature of the proposed test method is that it is based upon the use of monocytoid cells of human origin. It therefore by definition resembles more closely the actual response of humans. The two other test methods make use of either crustaceans (BET) or rabbits, both species more or less distinct from the human species. The response of humans, horseshoe crabs and rabbits toward Gram-negative endotoxin has been

studied extensively and the methods appear equivalent for this particular pyrogen (Cooper et al 1971; Greisman and Hornick, 1969). However, there are documented cases of medicinal products and specified pyrogenic substances that yield false-positive or false-negative results in either test method. Since the proposed test method is based on human cells, it may therefore predict more accurately the pyrogenicity of such substances in humans.

1.3.3 *Describe the intended range of substances amenable to the proposed test method and/or the limits of the proposed test method according to chemical class or physicochemical factors.*

The proposed test method is intended for the assessment of pyrogens in all parenteral medicinal products for human use, chemical or biological and including raw materials, bulk ingredients and excipients. Use of the proposed test method in testing environmental samples or medicinal products is suggested and may be feasible, but substantiating data are as yet limited or absent.

2 Test Method Protocol Components

2.1 Overview of test method.

Provide an overview of how the proposed test method is conducted. If appropriate, this would include the extent to which the protocol for the proposed test method adheres to established performance standards.

A highly detailed protocol describing the proposed test method (*Detailed protocol CRYO WB/IL-1: Human Whole Blood Pyrogen Test in 96-well Plates Using Fresh or Cryopreserved Blood; electronic file name: SOP CRYO WB IL 1*) is attached in Appendix A of this Background Review Document (BRD). It covers three variations to the preparation of the whole blood described under point 7: 7A - fresh blood using 96-well plates, 7B -cryopreserved blood according to the so called "PEI (Paul-Ehrlich-Institute, Langen, Germany) method" and 7C - cryopreserved blood according to the so called "Konstanz method".
The present BRD refers to the variation 7B; whereas theresults with variation 7A are included in Section 13 of BRD WB/IL-1 and variation 7C in Section 13 of this BRD.

The CRYO WB/IL-1 test method is a two-part assay for the detection of pyrogenic contamination. The test protocol itself can be divided into the following two parts:

1. Incubation of the sample with (diluted) human blood.
2. An enzyme linked immunoassay (ELISA) for the measurement of IL-1β.

Ad 1.
Diluted human whole blood is incubated overnight (10-24 hours) together with saline or RPMI and the sample of interest in sterile and pyrogen-free microtiter plate and aliquots are taken for further examination.

Ad 2.
Samples (aliquots of blood stimulation) are distributed into the wells of a microtiterplate which are coated with monoclonal antibodies specific for IL-1β.
An enzyme-conjugated polyclonal antibody against IL-1β is added. During a 90-minute incubation, a sandwich complex consisting of two antibodies and the IL-1β is formed. Unbound material is removed by a wash step.
A chromogenic substrate (3,3´,5,5´ -tetramethylbenzidine, TMB) reactive with the enzyme label is added. Color development is terminated by adding a stop solution after 30 minutes. The resulting color, read at 450 nm, is directly related to the IL-1β concentration. Bi-chromatic measurement with a 600-690 nm reference filter is recommended.

The WHO-LPS standard (code 94/580, E.coli O113:H10:K-), was used throughout the validation. This standard is identical to USP Reference Standard Endotoxin (EC6).
There are several possibilities to estimate the pyrogenic contamination of the preparations under test: 1) A quantitative estimation can be achieved by the construction of a dose-response curve for endotoxin standard (e.g. 5.0, 2.5, 1.0, 0.5 and 0.25 EU/ml) versus

optical density (OD) value of the IL-1β ELISA. The contamination of the preparations is expressed in endotoxin–equivalent units. 2) A qualitative test can be achieved by the inclusion of an endotoxin threshold control (e.g. one fixed dilution of the standard curve) which allows for the classification in positive and negative samples (i.e. pyrogenic and non-pyrogenic samples). 3) A qualitative test can also be achieved by inclusion of an appropriate positive product control.
A detailed description of analysis methods used during the validation of the test-method can be found in section 5 of the current BRD.

2.2 Rational for selected test components

Provide a detailed description and rationale, if appropriate, for the following aspects of the proposed test method:
2.2.1 Materials, equipment, and supplies needed.
The materials, equipment and supplies used for the CRYO WB/IL-1 test method are laboratory items that will be already available in a routine QC laboratory. There is no need for sophisticated or dedicated laboratory equipment throughout the test.
For all steps in the procedure, excluding the ELISA procedure, the materials (e.g. tips, containers, solutions) which will be in close contact with samples and blood cells need to be sterile and pyrogen free. The materials, equipment and supplies are specified in the protocol given in Appendix A. It should be realized that equivalent devices may also be used and it is the user's responsibility to validate the equivalence.

Materials for part 1: Blood Incubation with cryopreserved blood
Equipment
- Incubator (37°C + 5%CO_2)
- Multipette or adjustable 20 to 100 µl pipetters
- Multicannel pipettor 8 or 12 channels)
- Vortex mixer
- Laminar flow bench (recommended)

Consumables
- Non-pyrogenic 96-well polystyrene tissue culture microtiter plate (Falcon, 353072)
- Sterile and pyrogen-fee tips 20 and 100 µl.
- Combitips for multipette, 1.0 ml and 0.5 ml
- Reservoir for RPMI and saline
- Non-pyrogenic test tubes, preferably 12 or 15 ml centrifuge tubes (Greiner)

Materials for part 2: ELISA procedure
Equipment
- Multichannel pipettor
- Microplate mixer
- Microplate washer
- Microplate reader capable of readings at 450 nm (optional reference filter in the range of 600-690 nm)
- A software package facilitating data generation, analysis, reporting, and quality control

Consumables

- Graduated cylinder and plastic storage container for Buffered Wash Solution
- Tip-Tubs for reagent aspiration with Multichannel pipettor
- Non-sterile pitpette tips
- Non-sterile deionized water
- The IL-1β-ELISA kit (commercially obtained), containing:
  - IL-1 antibody coated micoplates.
  - Enzyme labeled antibody. Horseradish peroxidase-labeled, affinity-purified, polyclonal(rabbit) anti-IL-1β antibodies.
  - Endotoxin control.
  - Saline
  - TMB/Substrate solution
  - Buffered Wash Solution Concentrate (saline solution, with surfactants and preservative)
  - Stop Solution (acidic solution)

2.2.2 Dose-selection procedures, including the need for any dose range-finding studies or acute toxicity data prior to conducting a study, if applicable.
For every kind of test compound the interference with human blood and the Il-1β ELISA kit is determined. For this purpose, a preliminary "dose finding" test is conducted to establish a suitable (interference free) dilution for every new test compound. For the validation study (as described in section 4 of this BRD), the tested products were diluted according to their known ELC, which was usually far beyond interfering concentrations. The ELCs of the tested products or drugs were calculated according to the European Pharmacopoeia.

2.2.3 Endpoint(s) measured.
The proposed test method is an *in vitro* model of the fever response mechanism. It determines the release of interleukin-1β (IL-1β) by monocytoid cells present in human blood. IL-1β is released into the culture medium upon the interaction of pyrogens and specific receptors on the monocytoid cells. The measured endpoint IL-1β is one of the cytokines that trigger the fever response *in vivo*.

2.2.4 Duration of exposure.
The cryopreserved human whole blood is exposed to possible pyrogenic components in samples overnight (10-24 hours) in an incubator at 37°C + 5% CO_2. The monocytoid cells produce endogenous pyrogens if triggered by pyrogens. Before transferring the stimulation aliquots onto the IL-1β ELISA plate, the contents of the wells are thoroughly mixed.

2.2.5 Known limits of use.
The CRYO WB/IL-1 method described in the method protocol is not a finalized test system for the testing of all medicinal products. The method may be applied only to preparations that have been validated with this method, i.e. shown not to interfere with the blood and the IL-1β readout system at a specified dilution of the preparation. A paragraph describing the interference testing is included in the protocol (see Appendix

A). However, at this moment there are no medicinal products known that cannot be tested with the method.

2.2.6 Nature of the response assessed.
The proposed test method is an *in vitro* model of the fever response mechanism. Upon the interaction of exogenous pyrogens and specific receptors on the monocytoid cells endogenous pyrogens (e.g. interleukins, TNF-α and prostaglandins) are produced. In the body the fever response is triggered by these endogenous pyrogens. Immunoreactive IL-1β, the measured endpoint for the current method, is one of these endogenous pyrogens.

2.2.7 *Appropriate vehicle, positive, and negative controls and the basis for their selection.*
Throughout the development and validation phase the test compounds are diluted in 0.9% (w/v) clinical saline. This 0.9% clinical saline is considered an appropriate vehicle as no interference with active substances of a drug is to be expected.
In addition the test includes several controls.
A negative control: 0.9% clinical saline (sodium chloride)
A positive control: WHO-LPS 94/580, 0.5 EU/ml in clinical saline.
A negative product control: clean, released batch for each drug.
A positive product control: test item spike with WHO-LPS (code 94/580) at 0.5 EU/ml
The positive and negative controls are the same in every assay and are needed to establish the sensitivity of the test system. In addition, a product-based set of controls is used to reveal product-related interference.

2.2.8 Acceptable range of vehicle, positive and negative control responses and the basis for the acceptable ranges.
A CRYO WB/IL-1 assay is considered acceptable for further analysis if the mean OD value of the positive control (0.5 EU/ml) exhibits an OD that is greater than 1.6 times the mean OD over the negative control (0.9% clinical saline). The mean OD of the PPC is at 1.6 times the mean OD of the NPC or greater. The mean OD of the PPC has to be in the 50-200% range of the mean OD of the 0.5 EU/ml endotoxin control. The mean OD of the negative saline control is at 0.1 OD or lower. Moreover the response to different concentrations of the positive control should show a dose response relationship. To be able to quantify the responses to the positive control this should be well within the maximum response that can be measured with the test system.

As regards the substances to be tested, for products with an established ELC, specified in EU/ml, the product is diluted to its maximum valid dilution (MVD). The negative product control should be negative at the MVD. The response to the positive product control should be between 50% and 200% of the response to the positive control, indicating a possible pyrogenicity can be detected using these conditions.

2.2.9 Nature of the data to be collected and the methods used for data collection.
The raw data collected are the read-outs (absorbance) of the IL-1β ELISA, measured by an automated laboratory ELISA-plate reader. The wavelength is dependent on the chromogenic substrate applied, but when using TMB, the ELISA-plate is read at a

wavelength of 450 nm. Bi-chromatic measurement with a reference wavelength of 600-690 nm is recommended.

2.2.10 Type of media in which data are stored.
Data are stored in electronic files (windows98 compatible software) and as hard copy.

2.2.11 Measures of variability.
As part of the development of the WB/IL-1 test method (using fresh blood, see BRD WB/IL-1) the intralaboratory repeatability was assessed by independent and identical replicated measurement of the different concentrations of WHO-LPS. Furthermore, the limit of detection and its dependence from known but uncontrollable variables such us operator and blood donor were investigated. These variables and the inherent variation of biological systems make up to the total variation of the method. However, this part of the development was performed with fresh whole blood samples in reaction tubes. This part of the analysis of the variability was not repeated for the CRYO WB/IL-1 test method in 96-wells microtiter plates. However, while tested drugs spiked with LPS it is shown that the variability is comparable (one blinded experiment performed in three laboratories).

2.2.12 Statistical or nonstatistical methods used to analyze the resulting data, including methods to analyze for a dose-response relationship. Justify and describe the method(s) employed.
All experiments are run with four replicates of the test compound with cryopreserved blood from one batch on one plate. A standard curve in quadruplicate, using the International Standard for Endotoxin (calibrated in EU) is included. The assay should fulfill the minimum assay suitability requirements as detailed in the SOP. This includes the following criteria: The mean OD of the 0.5 EU/ml endotoxin control is at 1.6 times the mean OD of the negative saline control or greater. The mean OD of the PPC is at 1.6 times the mean OD of the NPC or greater. The mean OD of the PPC has to be in the 50-200% range of the mean OD of the 0.5 EU/ml endotoxin control.
The further analysis of the data was performed according to the procedure described in section 5.3 of the current CRYO WB/IL-1 test method.

2.2.13 Decision criteria and the basis for the prediction model used to classify a test chemical (e.g., positive, negative, or equivocal), as appropriate.
A prediction model (PM) was developed in order to classify substances as “pyrogenic for humans” or “non-pyrogenic for humans”. To be able to define a dichotome result in the alternative pyrogen test, a threshold pyrogen value of 0.5 EU/ml was chosen. This threshold value was based on historical data with rabbits (described in section 4.1). The suitability of the PM was assessed by testing substances which were artificially contaminated with endotoxin (substances are described in section 3.2 and 3.3).
The statistical approach, including quality criteria, is detailed in section 5.3

2.2.14 Information and data that will be included in the study report and availability of standard forms for data collection and submission.
Raw data were collected using a standard form. These were submitted to the quality department of ECVAM.

2.3 Basis for selection of this test method

Explain the basis for selection of the test method system. If an animal model is being used, this should include the rationale for selecting the species, strain or stock, sex, acceptable age range, diet, and other applicable parameters.

In view of the shortcomings of the rabbit pyrogen test and the BET test, *in vitro* pyrogen tests that utilize the exquisite sensitivity to exogenous pyrogen of monocytoid cells have been proposed. In such tests, products are incubated with human cell and the conditioned media assayed for pyrogenic cytokines (Duff & Atkins, 1982; Dinarello et al., 1984; Poole et al., 1988; Poole et al, 1989; Hansen and Christensen, 1990; Taktak et al., 1991; Bleeker et al., 1994).

The human whole blood assay was developed as a real *in vitro* alternative to the rabbit pyrogen test. The basic idea was to mimic the fever reaction in humans. In general, the detection of exogenous pyrogens (e.g. endotoxin) by blood cells causes them to release endogenous pyrogens like IL-1, IL-6 and TNF □. These cytokines affect the thermal regulation centre in the brain and increase the body temperature by changing its set point. In the past, several test methods have been developed that use the sensitivity of human peripheral blood monocytes to exogenous pyrogens. In an attempt to increase the sensitivity of these tests the monocytes/leukocytes were isolated from whole blood. In addition, various cell lines, which retain monocytoid characteristics, including the capacity to synthesize and secrete pyrogenic cytokines, have been studied.

However, the isolation of monocytes/leukocytes from whole blood as well as the maintenance of a cell-line is labour–intensive and time–consuming, technically sophisticated and requires expensive reagents. It is clear that using whole blood implies considerably simplified handling and that costs are limited. In an early stage of development of the assay, interleukin-1□ was most promising as the endogenous pyrogen used as the readout. In addition, a standardised version of the test in form of an interleukin-1□ kit is commercially available.

Fresh whole blood samples must be used with 4 hours after collection, which put logistic constrains on the WB/IL-1 assay. The method is limited by the availability of freshly drawn blood, putative safety concerns in the case of infected donors and interindividual donor differences. To overcome these limitations a method was developed and optimized to produce batches of cryopreserved blood that can be used directly after thawing without any washing steps. Applying cryopreserved blood (stored at -80^{o}C or liquid nitrogen) considerably improves the flexibility of the WB/IL-1 test method.

2.4 Proprietary components

If the test method employs proprietary components, describe what procedures are used to ensure their integrity (in terms of reliability and accuracy) from "lot-to-lot" and over time. Also describe procedures that the user may employ to verify the integrity of the proprietary components.

T. Hartung and A. Wendel are named as inventors in 436518000 (USPTO) 04/28/05 - Test procedure with biological system - Preparations containing deep-frozen blood are

used for test procedures for determining blood response.

It is stated in the method protocol that components supplied in the IL-1□ - ELISA kit are not interchangeable with other lots of the same components. Including the appropriate positive and negative controls in each run ensures the reliability and accuracy of the CRYO WB/IL-1 test method. As a positive control a specified amount of the Endotoxin Standard is used. Minimum assay suitability requirements are set and are described in the SOP (also summarized in section 2.2.12 of this BRD).

2.5 Replicates

Describe the basis for the number of replicate and repeat experiments; provide the rationale if experiments are not replicated or repeated.
All experiments are run with four replicates of the test compound on one plate. Outliers are rejected only after checking according to the Grubbs test ($p>0.05$). Four replicates is considered the minimal amount for the Grubbs test.

During the development phase of the WB/IL-1 test method, the intralaboratory reproducibility as well as the interlaboratory reproducibility of the WB/IL-1 test method was established by applying repeated experiments (described in WB/IL-1 BRD). The test method reliability (repeatability /reproducibility) was shown to be satisfactory while using freshly drawn blood. As only the fresh blood was replaced by cryopreserved blood it was considered feasible to accept this result for the CRYO WB/IL-1 as well. Therefore the accuracy of the CRYO WB/IL-1 while testing pharmaceutical substances (detailed in table 3.3.1) was established by performing single experiment in three participating laboratories (described in section 6 and 7). It is shown that accuracy of the CRYO WB/IL-1 method applying cryopreserved blood is indeed comparable with the WB/IL-1 applying freshly drawn blood. Appendix D includes data on the intralaboratory reproducibility of CRYO WB/IL-1 and KN CRYO WB/IL-1 (Section 13) during their development phase.

2.6 Modifications applied after validation

Discuss the basis for any modifications to the proposed test method protocol that were made based on results from validation studies.
The proposed test method protocol has not been modified during or after validation.

2.7 Differences with similar test methods

If applicable, discuss any differences between the protocol for the proposed test method and that for a comparable validated test method with established performance standards.
Not applicable.

3 Substances Used for Validation

3.1 Selection of substances used

Describe the rationale for the chemicals or products selected for use in the validation process. Include information on the suitability of the substances selected for testing, indicating any chemicals that were found to be unsuitable.

Selected test items were medicinal products available on the market. Released clinical batches were considered clean, i.e. containing no detectable pyrogens. To test the specificity, sensitivity and the interlaboratory reproducibility of the proposed test method, the products were spiked with pyrogen. For the present studies endotoxin (LPS) was selected as the model pyrogen, since it is well defined, standardized and readily available.

For the sensitivity and specificity the test items were assessed at their MVD. The MVD is the quotient of the ELC and the detection limit. The European Pharmacopoeia prescribes for various types of parenterals the amount of endotoxin that is maximally allowed in a medicinal product, i.e. the ELC, taking into consideration the dose, the route of administration and the dosing regimen of the product.

The aim of the study was to discriminate between negative and positive samples. Based on the selected pyrogen threshold value (see 4.1) and the PM applied, positive samples were defined as containing 0.5 EU/ml or more. Hence, to determine the MVD, the value of 0.5 EU/ml was defined as the detection limit.

Test items were assessed as such (negative product control), spiked with endotoxin at 0.5 IU/ml (positive product control) and after spiking with endotoxin at 5 levels (blinded samples). In addition a negative control (saline) and positive control (0.5 EU/ml in saline) were included to establish assay validity.

For interlaboratory reproducibility, the test items were tested, at a predefined dilution above the MVD, independently in 3 different laboratories, 3 times each. Based on the selected pyrogen threshold value (see 4.1) and the PM applied, positive samples were defined as containing 0.5 EU/ml or more. The test items were tested after spiking with endotoxin at four levels. For no other reasons but practical ones, i.e. availability of test materials, different test items were selected for this part of the validation study.

It was determined earlier whether candidate test items interfered with the outcome of the proposed test method. Interference was considered when the response of endotoxin in the diluted test item was below 50% or above 200% of the response of endotoxin in saline (spike-recovery). It was shown that none of the test items interfered with the assay at the selected dilutions (data not shown).

3.2 Number of substances

Discuss the rationale for the number of substances that were tested.

A total of 13 test items were selected for the validation study (see 3.3): 10 test items for determining sensitivity and specificity (table 3.3.1), 3 different test items for determining interlaboratory reproducibility (table 3.3.2). Test items and their spikes were

appropriately blinded by ECVAM before distribution to the participating testing facilities.

For sensitivity and specificity, each test item was tested after spiking at its individual MVD. Hence they each came with their own specific set of 5 endotoxin spike solutions: 0.0, 0.25, 0.5, 0.5 and 1.0 EU/ml. Simple logistics limited the amount of test items for this part of the validation study to 10. Since test items were assessed with 5 different endotoxin levels at 3 independent test facilities, this yielded a total of 150 data points, biometrically considered to be sufficient for further analysis.

For reproducibility each test item was spiked at 4 different levels (0.0, 0.0, 0.5 and 1.0 EU/ml) and tested at specified dilutions, 3 times at 3 laboratories.

3.3 Description of substances used

Table 3.3.1: Test items (parenteral drugs) used for determining sensitivity and specificity

Drug	code	Source	Agent	Indication	MVD (-fold)
Glucose 5% (w/v)	GL	Eifel	Glucose	nutrition	70
Ethanol 13% (w/w)	ET	B.Braun	Ethanol	diluent	35
MCP®	ME	Hexal	Metoclopramid	antiemetic	350
Orasthin®	OR	Aventis	Oxytocin	initiation of delivery	700
Binotal®	BI	Aventis	Ampicillin	antibiotic	140
Fenistil®	FE	Novartis	Dimetindenmaleat	antiallergic	175
Sostril®	SO	GlaxoSmithKline	Ranitidine	antiacidic	140
Beloc®	BE	Astra Zeneca	Metoprolol tartrate	heart dysfunction	140
Drug A*	LO	-	0.9% NaCl	-	35
Drug B*	MO	-	0.9% NaCl	-	70

*Drugs A and B were included as saline controls using notional ELCs.
Negative control: 0.9% clinical stock saline solution.
Positive control: WHO-LPS 94/580 (*E. coli* 0113:H10:K-), 0.5 EU/ml in clinical stock saline.

Table 3.3.2: Test items (parenteral drugs) used for determining reproducibility.

Drug	Source	Agent	Indication
Gelafundin®	Braun melsungen	Gelatin	Transfusion
Jonosteril ®	Fresenius	Electrolytes	Infusion
Haemate ®	Aventis	Factor VIII	Hemophilia

Negative control: 0.9% clinical stock saline solution.
Positive control: WHO-LPS 94/580 (*E. coli* 0113:H10:K-), 0.5 EU/ml in clinical stock saline.

BRD: CRYO WB/IL-1 March, 2006

3.4 Sample coding procedure

Describe the coding procedures used in the validation studies.

All test items are registered medicinal products and were obtained from a pharmaceutical supplier. Test items and endotoxin spiking samples were prepared, blinded where appropriate and coded under GLP. The blinding was performed by the quality assurance of ECVAM (Ispra, Italy). The Steinbein-Center InPut (University of Konstanz, Germany) was responsible for the preparation of samples and spikes and the subsequent shipment to each of the appropriate test facilities participating in the study.
For the sensitivity and specificity part of this study, test items and their respective spikes (5 per test item) were all blinded. For reproducibility testing, only the spikes (4) were blinded, the test items were not.

3.5 Recommended reference chemicals

For proposed test methods that are mechanistically and functionally similar to a validated test method with established performance standards, discuss the extent to which the recommended reference chemicals were tested in the proposed test method. In situations where a listed reference chemical was unavailable, the criteria used to select a replacement chemical should be described. To the extent possible, when compared to the original reference chemical, the replacement chemical should be from the same chemical/product class and produce similar effects in the in vivo reference test method. In addition, if applicable, the replacement chemical should have been tested in the mechanistically and functionally similar validated test method. If applicable, the rationale for adding additional chemicals and the adequacy of data from the in vivo reference test method or the species of interest should be provided.

The reference pyrogen material used was the international endotoxin standard WHO-LPS 94/580 (*E. coli* 0113:H10:K-). Where appropriate, the material was diluted in clinical saline solution (0.9%(w/v) sodium chloride). The saline was also used as negative control (blank).

4 *In vivo* Reference Data on Accuracy

4.1 Test protocol *in vivo* reference test method.

Provide a clear description of the protocol(s) used to generate data from the in vivo reference test method. If a specific guideline has been followed, it should be provided. Any deviations should be indicated, including the rationale for the deviation.

For ethical reasons, no rabbit pyrogen tests were performed for this study. However, Dr. U. Lüderitz-Püchel, Paul-Ehrlich Institute, Germany, kindly provided historical data, accumulated over several years, from 171 rabbits (Chinchilla Bastards). The respective pharmacopoeiae do not prescribe a rabbit strain for the *in vivo* pyrogen test, but Chinchilla rabbits are reported as a relatively sensitive strain for pyrogen testing.

The rabbits were injected with endotoxin and their rise in body temperature over the next 180 minutes was recorded (figure 4.1.1). From these data it was established that 50% of the rabbits got fever when treated with endotoxin at 5 EU/kg (Hoffmann et al, 2005a). Fever in rabbits is defined as a rise in body temperature over 0.55°C. On the basis of these historical animal data and corrected for the maximal volume allowed in rabbits, i.e. 10 ml per animal, a pyrogen threshold value of 0.5 EU/ml was defined for the PM in the proposed test method.

4.2 Accuracy

Provide the in vivo reference test method data used to assess the accuracy of the proposed test method. Individual human and/or animal reference test data, if available, should be provided. Provide the source of the reference data, including the literature citation for published data, or the laboratory study director and year generated for unpublished data.

As mentioned, no animal studies were done for ethical reasons. However, a theoretical assumption on specificity and sensitivity can be made (Hoffmann et al, 2005a). Taking the prevalence of the different final concentrations of LPS in the 5 spikes into account (1.0 EU/ml: 20%; 0.5 EU/ml: 40%; 0.25 EU/ml: 20% and 0.0 EU/ml: 20%) and calculating probabilities of misclassification using the chosen study design and defined threshold of pyrogenicity, i.e. 0.5 EU/ml, the theoretical sensitivity of the rabbit pyrogen test is 57.9% and the theoretical specificity of the rabbit pyrogen test is 88.3%.

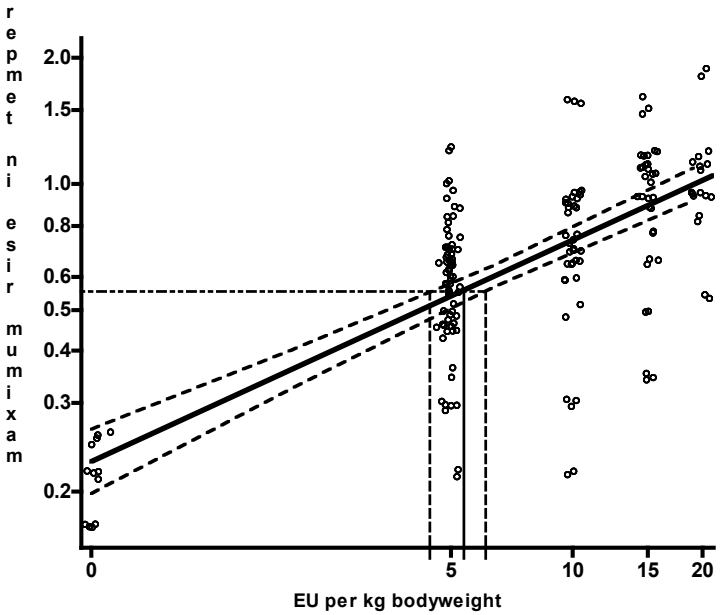


Figure 4.1.1 Dose-temperature of standard endotoxin applied to Chinchilla Bastards (n=171). Rabbits were treated with 1 ml saline containing 0, 5, 10, 15 and 20 EU of *E. coli* LPS (WHO-LPS 94/580 (E.coli O113:H10:K)) and their body temperature was measured over 180 min. Linear regression analysis was performed after logarithmic transformation of the data. Data are shown as dots to which a jitter-effect was applied in order to be able to distinguish congruent data. The full line depicts the linear regression whereas the dashed lines represent the 95%-confidence bounds. Furthermore, a horizontal line for a 0.55°C raise of temperature is added which is often defined as the rabbit threshold for fever. At the interception point of this line and the regression line 50% of the rabbits are to be expected to develop fever.

4.3 Original records

If not included in the submission, indicate if original records are available for the in vivo reference test method data.

The recognition of pyrogenic substances as bacterial by-products and the identification of a variety of pyrogenic agents enabled the development of a proper test to demonstrate non-pyrogenicity of the pharmaceutical product. As early as the 1920s, studies were done to select the most appropriate animal model. Results indicated that most mammals had a pyrogenic response, but only a few, including rabbits, dogs, cats, monkeys and horses showed a response similar to that in humans. For practical reasons, other species but rabbits and dogs were considered not practical. In 1942, Co Tui & Schrift described that rabbits are less thermo-stable as compared to dogs. Hence, rabbits are more suited for the purpose of testing for the absence of pyrogens, since a negative result is more significant.

4.4 Quality of data

Indicate the quality of the in vivo reference test method data, including the extent of GLP compliance and any use of coded chemicals.

Documented procedures were employed that were GLP-concordant. These were quality assured by quality assurance officers from ECVAM.

4.5 Toxicology

Discuss the availability and use of relevant toxicity information from the species of interest (e.g., human studies and reported toxicity from accidental or occupational exposure for human health-related toxicity testing).

Over time, a number of studies were done to correlate the rabbit test to pyrogenic reactions in humans. A conclusive study by Greisman and Hornick, published in 1969, who compared three purified endotoxin preparations (*Salmonella typhosa*, *E. Coli* and *Pseudomonas*) in New Zealand rabbits and in male volunteers, showed that the induction of a threshold pyrogenic response, on a weight basis, was similar to rabbit and man. At higher doses, rabbits respond less severe as compared to man.

4.6 Background on assay performance

Discuss what is known or not known about the accuracy and reliability of the in vivo reference test method.

As mentioned, animal studies were not performed for ethical reasons. However, a theoretical assumption on specificity and sensitivity can be made (Hoffmann et al, 2005a).Taking the prevalence of the different final concentrations of LPS in the 5 spikes into account (1.0 EU/ml: 20%; 0.5 EU/ml: 40%; 0.25 EU/ml: 20% and 0.0 EU/ml: 20%) and calculating probabilities of misclassification using the chosen study design and defined threshold of pyrogenicity, i.e. 0.5 EU/ml, the theoretical sensitivity of the rabbit pyrogen test is 57.9% and the theoretical specificity of the rabbit pyrogen test is 88.3%.

5 Test Method Data and Results

5.1 Test method protocol

Describe the proposed test method protocol used to generate each submitted set of data. Any differences from the proposed test method protocol should be described, and a rationale or explanation for the difference provided. Any protocol modifications made during the development process and their impact should be clearly stated for each data set.

The detailed protocol used during the catch-up validation of the CRYO WB/IL-1 test is provided in Appendix A of this BRD. It covers the precise step-by-step description of the test method and lists the necessary reagents and laboratory procedures for generating data. For the sake of clarity the protocol contains a detailed description of spiking with WHO-LPS and the dilution of the samples applied during the catch-up validation. A summary of the sample preparation is presented in table 5.1.2. The analysis described in section 5.3 of this BRD was applied to the data produced during catch-up validation.

The CRYO WB/IL-1 method was transferred from the developing laboratory (DL) to two other laboratories (denoted as naive laboratory 1 [NL1] and naive laboratory 2 [NL2]). All laboratories had to meet the validity criteria (minimum assay suitability requirements) as laid down in the SOP before the studies with medicinal substances were conducted.

The **interlaboratory reproducibility** was assessed by testing 3 different medicinal substances, Gelafundin, Jonosteril and Haemate (described in table 3.3.2, section 3.3.). Test items and their spikes were appropriately blinded. Test items were tested, at a predefined dilution above the MVD, independently in 3 laboratories. Test items were tested after spiking with WHO-LPS at four different levels, the spikes were blinded and coded by QA ECVAM. In addition a negative control (saline) and positive control (0.5 EU/ml) in saline were included to establish assay validity.
Although this part of the study was designed for assessment of interlaboratory reproducibility, a preliminary estimate of the accuracy could be derived from the data. Applying the PM to the results and evaluating the concordance in a two-by-two contingency table assessed accuracy.

To assess **accuracy** of the proposed test method 10 substances (listed in table 3.3.1), were spiked with five different concentrations of the WHO-LPS (one of which is negative). To permit the application of the chosen PM, each drug was diluted to its individual MVD, which has been calculated using the ELC to that drug (listed in table 3.3.1.). Each substance had their own specific set of endotoxin spike solution, ensuring that after spiking the undiluted drug, it contained 0.0; 0.25; 0.5; 0.5 and 1.0 EU/ml respectively when tested at its MVD. A detailed description of the sample preparation was supplied to the three independent laboratories (shown in table 5.1.2). To put more weight to the result of this part of the validation, the spikes were blinded and coded by QA ECVAM. The raw data of the CRYO WB/IL-1 assay are shown in paragraph 5.2. Accuracy was assessed by applying the PM to the results and evaluating the concordance in a two by two table. Intralaboratory reproducibility was (successfully) shown in

previous experiments with fresh whole blood samples and it is assumed that the variability was not affected by the change to cryopreserved blood assayed in 96-wells plates (see also Appendix D). Interlaboratory reproducibility was also assessed for the CRYO WB/IL-1 test method.

Table 5.1.2: Sample preparation for the testing of 10 substances spiked with 5 different concentrations of WHO-LPS.

<table>
<tr><td colspan="3">unblinded</td><td colspan="5">blinded</td></tr>
<tr><td colspan="3">dilution of drug up to MVD
↓</td><td colspan="5">spiking of undiluted drug: 0.5 ml each</td></tr>
<tr><td>diluted drug</td><td>NPC</td><td>PPC</td><td rowspan="2">+ 25 µl of Spike 1</td><td rowspan="2">+ 25 µl of Spike 2</td><td rowspan="2">+ 25 µl of Spike 3</td><td rowspan="2">+ 25 µl of Spike 4</td><td rowspan="2">+ 25 µl of Spike 5</td></tr>
<tr><td rowspan="2">0.5 ml</td><td>+ 25 µl saline</td><td>+ 25 µl PPC-LPS-spike *</td></tr>
<tr><td></td><td>(final conc. = 50 pg/ml)</td><td colspan="5">dilution to MVD
↓</td></tr>
<tr><td></td><td>test</td><td>test</td><td>test</td><td>test</td><td>test</td><td>test</td><td>test</td></tr>
</table>

** PPC-LPS-spike contains 1050 pg/ml = 21fold 50 pg/ml*

NPC = Negative Product Control, **PPC** = Positive Product Control, **MVD** = Maximal Valid Dilution

5.2 Accuracy and reliability

Provide all data obtained to evaluate the accuracy and reliability of the proposed test method. This should include copies of original data from individual animals and/or individual samples, as well as derived data. The laboratory's summary judgment regarding the outcome of each test should be provided. The submission should include data (and explanations) from all studies, whether successful or not.
See figures 5.2.1 and 5.2.2 for the prevalidation data. The data of the validation are presented as tabulated results in section 5.4.

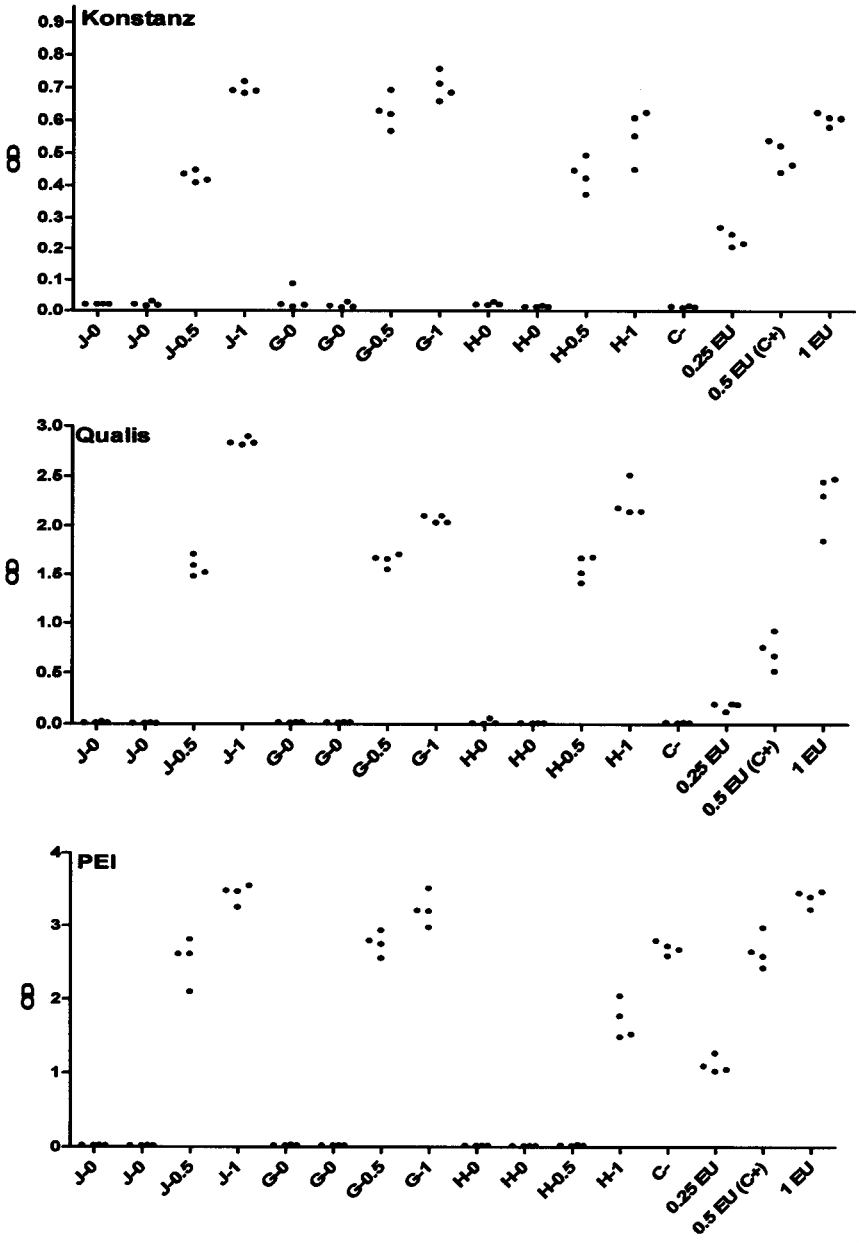


Figure. 5.2.1 Prevalidation data for the CRYO WB/IL-1 test method (Cryopreserved blood prepared by the PEI method). Three different drugs were spiked (blinded) with WHO-LPS at 0.0, 0.0, 0.5 and 1.0 IU/ml, respectively. Experiment was run independently at three different laboratories. Treatment and controls are abbreviated (J = Jonosteril; G = Gelafundin ; H = Heamate. C- = negative control (saline); C+ = positive control (0.5 IU/ml in saline).

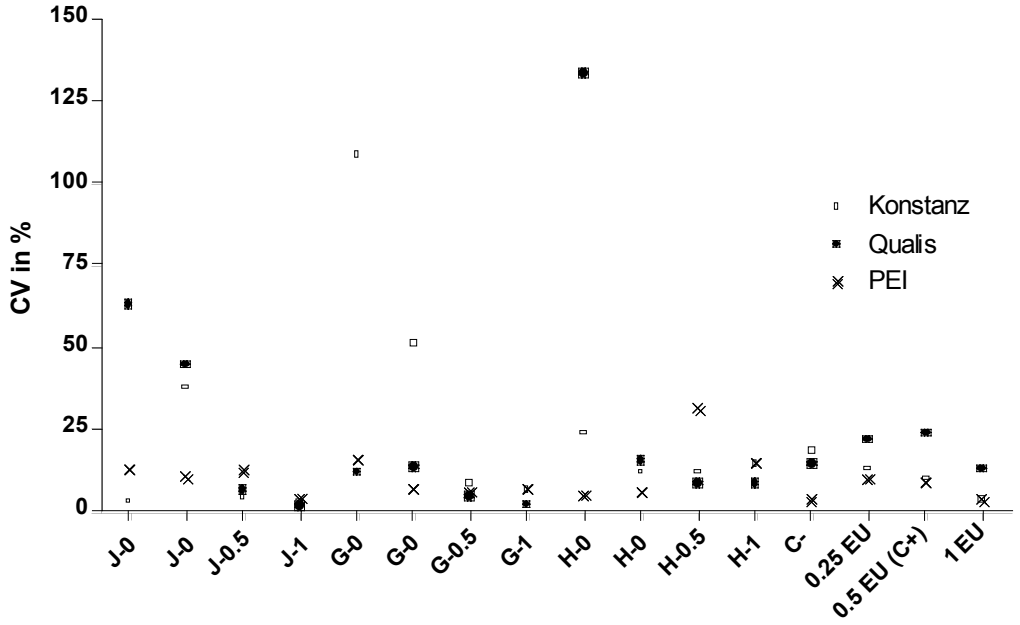


Figure. 5.2.2: Coefficient of variation (CV) of the prevalidation data (readout of IL-1□ ELISA) of CRYO WB/IL-1 test method. CV of the three different drugs spiked (with WHO-LPS at 0.0, 0.0, 0.5 and 1.0 IU/ml, respectively. Experiment was run independently at three different laboratories (Konstanz, Qualis and PEI).
Treatment and controls are abbreviated (J = Jonesteril; G = Gelafundin ; H = Heamate. C- = negative control (saline); C+ = positive control (0.5 IU/ml in saline).

5.3 Statistics

Describe the statistical approach used to evaluate the data resulting from studies conducted with the proposed test method.

A generally applicable analytical procedure was employed. This procedure includes a universal PM as well as quality criteria. First, a two-step procedure consisting of a variance-criterion and an outlier-test was applied. For this, the Dixon's test (Barnett and Lewis, 1984), which is USP approved, was chosen with the significance level of α=0.01 and applied to identify and eliminate aberrant data.

Next, the negative and the respective positive control are compared to ensure a suitable limit of detection. For this, a one-sided t-test with a significance level of α=0.01 is applied to the ln-transformed data to ensure that the response to the positive control is significantly larger than that of the respective negative control.

Finally, the samples are classified as either negative or positive by the outcome of a one-sided version of the t-test, which is based on the assigned pyrogen threshold value. The final results will be given in 2 x 2 contingency tables (table 5.3.1). These tables allow for estimation of accuracy (sensitivity and specificity) and reproducibility of the proposed test method.

Table 5.3.1: 2x2 contingency table.

		pre-defined class ("truth")		Σ
		1	***0***	
Classification by test system and PM	***1***	***a***	***b***	$a+b = n_{.1}$
	0	***c***	***d***	$c+d = n_{.0}$
Σ		$a+c = n_{1.}$	$b+d = n_{0.}$	***n***

Accuracy:

The most important statistical tool to determine accuracy (specificity and sensitivity) is the PM (Hothorn, 1995). In general, it is a statistical model, which classifies a given drug by an objective diagnostic or deciding rule. The objective of a dichotomous result requires a clear cut PM, which assigns a drug in one of the two classes "pyrogenic for humans" and "non-pyrogenic for humans". Since a threshold pyrogen value will be used, a one-sided test is appropriate for the task. Because the data are normalised by a ln-transformation, a t-test is chosen. Although the variances over the range of concentration converge by the transformation, the assumptions of equal variances do generally not hold true, because it depends on additional covariates. Therefore, the one sided Welch-t-test (Snedecor and Cochran, 1989) is applied. Due to the safety aspect of the basic problem, the hypotheses of the test are

$$H_0 : \mu_{S_i j} > \mu_{S+} \quad vs \quad H_1 : \mu_{S_i j} < \mu_{S+},$$

where $\mu_{...}$ denotes the parameter of location of the respective ln-transformed distribution. This approach controls the probability of false positive outcomes directly by means of its significance level α, which is chosen as 0.01, because is assumes hazard, respectively pyrogenicity, of the tested drug in H_0, and assures safety, i.e. non-pyrogenicity. The test statistic is

$$T_{S_i j} = \frac{\bar{x}_{S+} - \bar{x}_{S_i j}}{\sqrt{\frac{s^2_{S+}}{n_{S+}} + \frac{s^2_{S_i j}}{n_{S_i j}}}}.$$

The PM is built by means of the outcome of the test. Let 0 denote safety and 1 denote hazard. The classification of S_i-j is then determined by

$$S_ij = 0, \text{ if } \quad T_{S_i j} > t_{0.99; n_{S+} + n_{S_i j} - 2},$$

$$S_ij = 1, \text{ else,}$$

where $t_{0.99;n_{S+}+n_{S_ij}-2}$ the 0.99-quantile of the t-distribution with $n_{S+} + n_{S_ij} - 2$ degrees of freedom. The number of replicates for every control and sample, i.e. $n_{...}$, was harmonised to be four. Due to the possibility of removing one observation by the outlier test, the number of replicates could be reduced to three. The classification of a version of a drug is regarded as an independent decision. Therefore, the niveau α is local.

Finally, the classifications of the drugs will be summarised in 2x2 contingency table (table 3). From these tables, estimates of the sensitivity (S_E), i.e. the probability of correctly classified positive drugs and specificity (S_P), i.e. the probability of correctly classified negative drugs, will be obtained by the respective proportions. Where

$$S_E = a / (a + c) * 100\%$$
and
$$S_P = d / (b + d) * 100\%.$$

Furthermore, these estimates will be accompanied by confidence intervals, which will be calculated by the Pearson-Clopper method [15]. For example, let $\hat{p}_{SE}$ denote the proportion, namely the sensitivity, under investigation. Then the confidence interval to a niveau α is calculated as

$$\left[p_{SE}^{L} = \frac{\mathrm{a}F_{2\mathrm{a};2(n_{1\cdot}-\mathrm{a}+1);\frac{\alpha}{2}}}{n_{1\cdot} - \mathrm{a} + 1 + \mathrm{a}F_{2\mathrm{a};2(n_{1\cdot}-\mathrm{a}+1);\frac{\alpha}{2}}}; p_{SE}^{U} = \frac{(\mathrm{a}+1)F_{2(\mathrm{a}+1);2(n_{1\cdot}-\mathrm{a});1-\frac{\alpha}{2}}}{n_{1\cdot} - \mathrm{a} + (\mathrm{a}+1)F_{2(\mathrm{a}+1);2(n_{1\cdot}-\mathrm{a});1-\frac{\alpha}{2}}} \right],$$

where F denotes the respective quantile of the F-distribution and $n_{1\cdot}$ is the sample size of the positive drugs and a the number of correctly classified drugs.

By contaminating the drugs artificially and by defining a threshold value, which is assumed to be appropriate, the class of a drug is determined beforehand. The versions of drugs, which are effectively contaminated, but below the threshold dose, are considered to be negative, respectively safe, because their contamination is not crucial for humans in terms of ELC.

Reproducibility:

The analysis of the interlaboratory reproducibility was assessed from the three identical and independent runs conducted in each of 3 laboratories. The comparison of the three runs was carried out blindly such that the testing facility did not know the true classification of the sample, either pyrogenic or non-pyrogenic. By this procedure only the randomness and reproducibility of the methods was assessed and not systematic errors, which may have arisen from other sources, e.g. logistics or preparation of the samples. This means that if a sample was (mis)classified in all three runs the result is reproducible regardless of the (mis)classification of the sample. Therefore, a measure of similarity, i.e. complete simple matching with equal weights, was preferred to the coefficient of correlation for 2x2 contingency tables.

BRD: CRYO WB/IL-1 March, 2006

The study was designed as follows: each laboratory had to conduct three independent runs with the same 12 samples (3 test items with 4 blinded spikes each) and two controls, i.e. saline as a negative control (C-) and a 0.5 EU/ml LPS-spike in saline as a positive controls (C+). The samples were derived from the three substances Gelafundine, Haemate and Jonosteril. Per run, each substance was blindly spiked twice with saline, once with 0.5 EU/ml LPS and once with 1 EU/ml LPS, which resulted in a balanced design with regard to positive and negative samples, i.e. samples expected to be pyrogenic and non-pyrogenic, respectively.

The three independent runs per testing facility provide the information on which the assessment of the intralaboratory reproducibility is based. The combined results of the three runs per testing facility were used to determine interlaboratory reproducibility. The correlation of the prediction (in terms if the Bravais-Pearson coefficient of correlation) between all runs is calculated, independent of whether that classification is true or false. A BP-correlation of 1 is calculated, if two runs gave exactly the same predictions for the twelve substances. If one run gives adverse classifications for all substances than the other, the correlation is –1. As these calculations do not need information of the true status of a sample, they were carried out blinded.

5.4 Tabulated results

Provide a summary, in graphic or tabular form, of the results.
See tables 5.4.1 and 5.4.2.

Table 5.4.1: Results of prevalidation, testing of 3 spiked substances by the three involved laboratories. Classifications after applying the PM (compare to fig. 5.2.1)
"0"denotes "non-pyrogenic"; "1" denotes "pyrogenic".

drug	Code	spike in EU	laboratory		
			Konstanz	Qualis	PEI
Jonosteril	J - 0 (1)	0	0	0	0
	J - 0 (2)	0	0	0	0
	J - 0.5	0.5	1	1	1
	J - 1	1	1	1	1
Gelafundin	G - 0 (1)	0	0	0	0
	G - 0 (2)	0	0	0	0
	G - 0.5	0.5	1	1	1
	G - 1	1	1	1	1
Haemate	H - 0 (1)	0	0	0	0
	H - 0 (2)	0	0	0	0
	H - 0.5	0.5	1	1	**0**
	H - 1	1	1	1	1

Table 5.4.2: Results of the validation study of 10 drugs, spiked with WHO-LPS at 0.0, 0.25, 0.5, 0.5 and 1.0 IU/ml, respectively and tested in 3 different laboratories. Samples and spikes were blinded. Classifications after applying the PM (compare to fig. 5.2.7).

drug (code)	spike EU/ml	"truth"	results PEI	Qua	Nov
Beloc (BE)	0.00	0	0	0	CV
	0.25	0	0	1	CV
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Binotal (BI)	0.00	0	0	0	0
	0.25	0	0	1	1
	0.50	1	0	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Ethanol 13% (ET)	0.00	0	0	0	nq
	0.25	0	CV	1	nq
	0.50	1	1	1	nq
	0.50	1	1	1	nq
	1.00	1	1	1	nq
Fenistil (FE)	0.00	0	0	0	0
	0.25	0	0	1	1
	0.50	1	1	1	CV
	0.50	1	1	1	1
	1.00	1	1	1	1
Glucose 5% (GL)	0.00	0	0	0	0
	0.25	0	CV	0	0
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
"Drug A"	0.00	0	CV	0	nq
0.9% NaCl (LO)	0.25	0	0	0	nq
	0.50	1	1	1	nq
	0.50	1	1	1	nq
	1.00	1	1	1	nq
MCP (ME)	0.00	0	0	0	0
	0.25	0	CV	1	CV
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
"Drug B"	0.00	0	0	0	nq
0.9% NaCl (MO)	0.25	0	0	0	nq
	0.50	1	1	1	nq

drug (code)	spike EU/ml	"truth"	results PEI	Qua	Nov
	0.50	1	1	1	nq
	1.00	1	1	1	nq
Orasthin (OR)	0.00	0	0	0	0
/Syntocinon	0.25	0	0	CV	CV
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Sostril (SO)	0.00	0	0	0	nq
	0.25	0	0	**1**	nq
	0.50	1	**0**	1	nq
	0.50	1	1	1	nq
	1.00	1	1	1	nq

0 = considered/classified negative
1 = considered/classified positive

Grey shading indicates that for these drugs the PPCs did not qualify so that the PC was used in the PM. CV = sample showed a variability resulting in exclusion, i.e. CV > 45% and no significant outlier present.
nq = not qualified according to quality criteria, i.e. failure of PPCs and PCs
False classifications are in bold type.

5.5 Coding of data

For each set of data, indicate whether coded chemicals were tested, whether experiments were conducted without knowledge of the chemicals being tested, and the extent to which experiments followed GLP guidelines.
Blinding of drugs and/or spikes is indicated with the data.

5.6 Circumstances

Indicate the "lot-to-lot" consistency of the test substances, the time frame of the various studies, and the laboratory in which the study or studies were conducted. A coded designation for each laboratory is acceptable.
In each part of the study, all samples are derived from one (clinical) lot.

5.7 Other data available

Indicate the availability of any data not submitted for external audit, if requested.
All relevant data were submitted with the present BRD.

6 Test Method Accuracy

6.1 Accuracy

Describe the accuracy (e.g., concordance, sensitivity, specificity, positive and negative predictivity, false positive and negative rates) of the proposed test method compared with the reference test method. Explain how discordant results in the same or multiple laboratories from the proposed test were considered when calculating accuracy.
Test method accuracy was assessed in two large scale experiments performed with the drugs outlined in table 3.3.1 and table 3.3.2 in section 3 respectively. As described before, a prevalidation test was conducted with 3 different drugs and in the final validation experiment ten drugs were tested in the three participating laboratories. From the first experiment a preliminary estimate of sensitivity and specificity can be calculated, whereas the second is regarded as the established accuracy for the CRYO WB/IL-1 assay.

6.1.1 Preliminary estimate of the accuracy of the CRYO WB/IL-1 test. In the prevalidation phase of the study the developing laboratories (DLs) determined for each drug (outlined in table 3.3.2, section 3.3) the smallest dilution within the MVD that showed no interference or an acceptable degree of interference with the spike recovery. In general the lowest dilution of the sample allowing for a 50-200% spike recovery was chosen. In addition, the positive control (PC) set at 0.5 EU/ml saline was used as the classification threshold. The laboratory procedure as described in the method protocol was maintained throughout the study. Although it was realized there were some drawbacks to the concept for interference testing and applying the PC as a threshold, this small scale study allows for a preliminary estimate of the accuracy of CRYO WB/IL-1 method. It has to be noted that this part of the study was designed to provide an estimate of the interlaboratory reproducibility. Therefore it will also be discussed in detail in section 7 (Test Method Reliability).

According to the PM applied during an early phase of the study the outcome (positive/negative) is related to the positive control (PC=0.5 EU/ml). If absorbance of sample > absorbance of PC, then the sample is classified as positive. If absorbance of sample < PC, then the sample is classified as negative. While performing the experiments during this phase it was realized that there is a flaw between interference testing and the PM. After the interference testing, the lowest dilution of the sample allowing for a 50-200% spike recovery was chosen. It is obvious that even with a flawless repeatability of the assay; a spike recovery between 50%-100% would be classified as negative according to the preliminary PM. In addition, due to unforeseen problems with the preparations of the spike, the recovery of the spikes was far below 100%. (This is outside the scope of the study and will not be discussed). As a consequence of the employed preliminary setup of the study the sensitivity will be underestimated, and the specificity will be overestimated.

In short, three test substances were spiked with WHO-LPS at four levels (0, 0, 0.5 and 1.0 EU/ml respectively), which resulted in 12 samples with a balanced design with regard to positive and negative samples (i.e. samples expected to be pyrogenic and non-

pyrogenic respectively. These 12 samples were tested in three laboratories. In total there were 36 classifications in 3 laboratories. Results are described in detail in section 7. A 2x2 contingency table was constructed (table 6.1.1), from which the estimates of sensitivity and specificity can easily be derived.

Table 6.1.1: 2x2 contingency table. The prediction model applied to a preliminary validation study.

		True status of samples +	-	Total
PM	+	17	0	16
	-	1	18	20
	Total	18	18	36

The specifications of specificity and sensitivity described in section 5.3 were applied to these results. The specificity (Sp) of the CRYO WB/IL-1 assay is 100% (18/(18+0)*100%). The sensitivity (Se) calculated for this data set is 94.4% (17/(17+1) *100%). As outlined previously the specificity is overestimated and the sensitivity is underestimated as a result of the design of this part of the study.

6.1.2 Test method accuracy of the proposed CRYO WB/IL-1 method. To assess accuracy of the proposed method, 10 substances (listed in table 3.1.1, section 3) were spiked with five different concentrations of the WHO-LPS (one of which is negative). Thus, in total, 50 samples have been tested in each laboratory.
To permit the application of the chosen PM, each drug was diluted to its individual MVD, which has been calculated using the ELC to that drug (listed in section 3). Lesser dilutions were tested by the DL, and showed no interference. Therefore interference was not expected at the individual MVD. Each substance had their own specific set of endotoxin spike solution, ensuring that after spiking the undiluted drug, it contained 0.0; 0.25; 0.5; 0.5 and 1.0 EU/ml respectively when tested at its MVD. A detailed description of the sample preparation was supplied to the three independent laboratories (shown in table 5.1.1 for convenience). To put more weight to this part of the validation, the spikes were blinded and coded by QA ECVAM. Accuracy was assessed by applying the PM to the results (summarized in table 5.4.2) and evaluating the concordance in this section in a two by two contingency table (table 6.1.2 and 6.1.3). As described above 10 substances, spiked with 5 different WHO-LPS concentrations were tested in three laboratories and consequently a maximum of 150 data were available for analysis.

Of the 150 available data for the CRYO WB/IL-1 method, eleven sets of 4 replicates showed a high variability resulting in exclusion, i.e. CV > 45% and no significant outliers present. Therefore 139 data in total could be used to estimate the specificity and sensitivity of the CRYO WB/IL-1 method. The results are shown separately for each participating laboratory (table 6.1.2) as well as combined for all these laboratories (table 6.1.2).

BRD: CRYO WB/IL-1 March, 2006

The specificity that can be estimated from the available results for DL, NL1 and NL2 is 68.4%, 75% and 100% respectively The estimated sensitivity of the CRYO WB/IL-1 assay was excellent of all three participating laboratories: 93.3%, 100% en 100% respectively (calculated from results in table 6.1.2).

Table 6.1.2: 2x2 contingency table. Prediction model applied to the CRYO WB/IL-1 test result of 10 different substances assessed in three different laboratories. Results of each laboratory separately (DL, NL1 and NL2= PEI, Qualis and Novartis respectively).

Results DL		True status of samples +	-	Total
PM	+	28	0	28
	-	2	16	18
Total		30	16	46

Results NL1		True status of samples +	-	Total
PM	+	30	6	36
	-	0	13	13
Total		30	19	49

Results NL2		True status of samples +	-	Total
PM	+	17	2	19
	-	0	6	6
Total		17	8	25

The specificity of the combined results of the three laboratories of the assay is 81.4% (35/(35+8)*100%), 95% confidence interval [0.679-0.920]. The sensitivity equals 97.4% (75/(75+2) *100%), 95% confidence interval [0.907-0.997]. (Summarized in table 6.1.3 and 6.1.4).

Table 6.1.3: 2x2 contingency table. Prediction model applied to the CRYO WB/IL-1 test result of 10 different substances assessed in three different laboratories. Combined results.

		True status of samples +	True status of samples -	Total
PM	+	75	8	83
	-	2	35	37
	Total	77	43	120

Table 6.1.3: Specificity and sensitivity of the CRYO WB/IL-1 assay

	N total	N correctly identified	proportion	95% CI lower limit	95% CI upper limit
Specificity (Sp)	43	35	81.4%	0.679	0.920
Sensitivity (Se)	77	75	97.4%	0.907	0.997

6.2 Concordancy to *in vivo* reference method

Discuss results that are discordant with results from the in vivo reference method.
Not applicable.

6.3 Comparison with reference methods

Discuss the accuracy of the proposed test method compared to data or recognized toxicity from the species of interest (e.g., humans for human health-related toxicity testing), where such data or toxicity classification are available. This is essential when the method is measuring or predicting an endpoint for which there is no preexisting method. In instances where the proposed test method was discordant from the in vivo reference test method, describe the frequency of correct predictions of each test method compared to recognized toxicity information from the species of interest.
Not applicable.

6.4 Strength and limitations

State the strengths and limitations of the proposed test method, including those applicable to specific chemical classes or physical-chemical properties.
It appears the proposed test is applicable to most classes of medicinal products, at least those that are non- or low-toxic to cells *in vitro*. I addition, the test may be employed to assess pyrogenicity of various medical devices, such as (biological) bovine collagen bone implants.

6.5 Data interpretation

Describe the salient issues of data interpretation, including why specific parameters were selected for inclusion.
No issues.

BRD: CRYO WB/IL-1 March, 2006

6.6 Comparison to other methods

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the results obtained with both test methods should be compared with each other and with the in vivo reference test method and/or toxicity information from the species of interest.

Not applicable.

7 Test Method Reliability (Repeatability/Reproducibility)

7.1 Selection of substances

Discuss the selection rationale for the substances used to evaluate the reliability (intralaboratory repeatability and intra- and interlaboratory reproducibility) of the proposed test method as well as the extent to which the chosen set of substances represents the range of possible test outcomes.

The rationale for the selection of the substances is described in section 3.3. In short: for the present studies endotoxin (WHO-LPS) was selected as the model pyrogen, since it is well defined biological standard and readily available. Selected test substances were medicinal products available on the market. These batches are released by the manufacturers and comply with the Marketing Authorisation file and European Pharmacopoeia. Therefore these batches are considered to contain no detectable pyrogens. To test the method reliability the medical products were spiked with endotoxin.

7.2 Results

Provide analyses and conclusions reached regarding the repeatability and reproducibility of the proposed test method. Acceptable methods of analyses might include those described in ASTM E691-92 (13) or by coefficient of variation analysis.

Interlaboratory reproducibility.

The interlaboratory reproducibility was assessed in a prevalidation test by testing 3 different medicinal substances, Gelafundin, Jonosteril and Haemate (described in table 3.3.2, section 3.3.). Test substances and their spikes were appropriately blinded. Test substances were tested, at a predefined dilution above the MVD, independently in 3 different laboratories, 3 times each. The three test substances were spiked with WHO-LPS at four levels (0, 0, 0.5 and 1.0 EU/ml respectively), which resulted in 12 samples with a balanced design with regard to positive and negative samples (i.e. samples expected to be pyrogenic and non-pyrogenic respectively. In addition a negative control (saline) and positive control (0.5 EU/ml in saline) were included to establish assay validity. To avoid interference, the DL performed interference testing in terms of the BET, i.e. 50-200% spike recovery, and decided on the dilution of the test substances. Dilutions chosen for Gelafundine, Haemate, Jonosteril were 1:2, 1:20 and 1:2 respectively.

For the CRYO WB/IL-1 test method, the coefficient of variation (CV) of the prevalidation test was assessed. The prevalidation test, three different drugs spiked with WHO-LPS (0.0, 0.0, 0.5 or 1.0 IU/ml), was run independently at three different laboratories (Konstanz, Qualis and PEI). The CVs were calculated for each treatment or control for all laboratories (shown in figure 5.2.2). While the major part of the CVs was smaller than 30%, four samples (Konstanz: both G-0; Qualis:J-0, H-0) showed a CV larger than 45%. In all of these, one replicate was much larger than the others, which gave very low responses (less than 0.02 OD).

Also for the fresh blood WB/IL-1 test, described in a separate BRD, it was concluded that the coefficient of variation for the sets of four replicates in the WB/IL-1 assay is usually below 45%. This is considered acceptable for a biological assay. To harmonize the acceptance criteria between different variations of the WB/IL-1 test, the coefficient of variation was arbitrarily set at CV<45% for the CRYO WB/IL-1 test.
The intralaboratory reproducibility of the CRYO WB/IL-1 test method is not assessed in this study (see Appendix D), but is considered to be similar to the WB/IL-1 test method (using fresh blood and conducted in test tubes) and therefore acceptable. As the interlaboratory reproducibility (usually worse than the intralaboratory reproducibility) is indeed shown to be satisfactory for the CRYO WB/IL-1 method presented in this BRD, this assumption was proven to be valid.
The analysis of the interlaboratory reproducibility could be assessed from the identical and independent runs conducted in the three laboratory. The three runs was carried out blindly such that the laboratory did not know the true classification of the sample (either pyrogenic or non-pyrogenic). By this procedure only the randomness and reproducibility of the methods was assessed and not systematic errors, which may have arisen from other sources, e.g. logistics or preparation of the samples. This means that if a sample was misclassified in all three laboratories the result is 100% interlaboratory reproducible (regardless of the misclassification of the sample).
According to the preliminary PM applied during this phase of the study the outcome (positive/negative) was related to the positive control (PC=0.5 EU/ml). If absorbance of sample > absorbance of PC, then sample is classified as being positive. If absorbance of sample < PC, sample is classified as negative (positive/pyrogenic = 1, negative/non-pyrogenic = 0). (NB. During this phase it was realized that there is a flaw between interference testing and the PM. After the interference testing, the lowest dilution of the sample allowing 50-200% spike recovery was chosen. It is obvious that even with a flawless repeatability of the assay, a spike recovery between 50%-100% would be classified as negative according to the preliminary PM.)
The interlaboratory reproducibility of the CRYO WB/IL-1 method was assessed by comparing the results of the preliminary test of the three laboratories. The measure of similarity is the proportion of equally classified samples. These proportions are summarized in table 7.2.3, and indicate that there is a good interlaboratory reproducibility for the CRYO WB/IL-1 test of at least 91.7%.

Table 7.2.3: Interlaboratory reproducibility assessed by interlaboratory correlations. Result of testing 3 substances (with four different spikes) by 3 laboratories.

Laboratories	**Interlaboratory reproducibility**	**Number of equal predictions**
DL – NL1	91.7%	11/12
DL – NL2	91.7%	11/12
NL1 – NL 2	91.7%	11/12
Mean	91.7%	
Same in three laboratories	83.3%	10/12

DL = Konstanz; NL1 = Qualis; NL2 = PEI

Also from the result of the large scale study (testing 10 substances spiked with 5 separate spikes), the interlaboratory reproducibility can be estimated (table 7.2.4). None of the laboratories identified all samples correctly. The reproducibility varied from 88.4% to 100% between two laboratories. From the available results it can be concluded that also the reproducibility between all three participating laboratories was satisfactory: All three laboratories found the same result for 21 out of 24 samples (87.5%).

Table 7.2.4: Interlaboratory reproducibility: Assessed by testing of 10 substances, spiked 5 times. One run of 50 samples by three different laboratories.

Laboratories	**Interlaboratory Reproducibility**	**Number of equal predictions**
DL - NL1	84.4%	38 / 45
DL - NL2	87.5%	21 / 24
NL1 – NL2	100%	25 / 25
Mean	90.6%	
same result in all laboratories	87.5%	21 / 24

DL = PEI; NL1 = Qualis; NL2 = Novartis.

Conclusion: The results of the prevalidation experiment indicated that the interlaboratory reproducibility was satisfactory. The reproducibility between two laboratories equals 91.7% during prevalidation. This was confirmed in the validation study where the interlaboratory reproducibility ranged from 88.4% to 100%. All three participating laboratories predicted the same in 83.3% and 87.5% of the measurements respectively.

7.3 Historical data

Summarize historical positive and negative control data, including number of experiments, measures of central tendency, and variability.
Not applicable.

7.4 Comparison to other methods

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the reliability of the two test methods should be compared and any differences discussed.
Not applicable.

8 Test Method Data Quality

8.1 Conformity

State the extent of adherence to national and international GLP guidelines (7-12) for all submitted data, including that for the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method. Information regarding the use of coded chemicals and coded testing should be included.
The studies were done in accordance to the guidelines for GLP. Written protocols and approved standard operating procedures were followed during the entire course of the study. Deviations were recorded and, where appropriate, approved in amendments. All data are stored and archived. As mentioned, samples were appropriately blinded.

8.2 Audits

Summarize the results of any data quality audits, if conducted.
No audits were conducted.

8.3 Deviations

Discuss the impact of deviations from GLP guidelines or any noncompliance detected in the data quality audits.
Not applicable.

8.4 Raw data

Address the availability of laboratory notebooks or other records for an independent audit. Unpublished data should be supported by laboratory notebooks.
All records are stored and archived by the contributing laboratories and available for inspection.

9 Other Scientific Reports and Reviews

9.1 Summary

Summarize all available and relevant data from other published or unpublished studies conducted using the proposed test method.

Relevant data obtained with the WB/IL-1 using fresh blood (see BRD WB/IL-1) are described in a number of published studies and reports. The most important ones for this BRD are included in the Appendix B as hardcopies and referenced in Section 12, whereas for others only the references are given in section 14. In most of the study reports the WB/IL-1 is named *in vitro* pyrogen test or IPT.

The establishment of the whole blood test as an alternative to the rabbit pyrogen test as well as the comparison to the BET is described below.

Further applications were developed by adaptation to the basic whole blood test e.g. to measuring pyrogenic contaminations of medical devices and measuring the air quality in the working place and references are included in Section 14 in part 2 and 3.

A total of 96 batches of parenteral pharmaceuticals from 21 indication groups were tested using the WB/IL-1 test and compared to data from the rabbit and BET test, if available (Jahnke et.al., 2000). For these batches of parenteral drugs it was shown that the results of the three methods correlate well. In one case (an amino acid-containing infusion solution) a pyrogen-containing batch was clearly detected by all three testing systems. The other parenteral pharmaceuticals remained negative in all assays. It is worth mentioning that all of the products could be tested with the WB/IL-1, in some cases after interfering factors had been excluded. A few drugs (e.g. dopamine) were found to affect the sensitivity of the WB/IL-1 and hence caused interference, but this could be overcome by diluting the drug.

In a preliminary study (Fennrich et al., 1999), the suitability of the WB/IL-1 was tested by determining the LPS retrieval in spiked pharmaceutical samples at the border line concentrations given in the European Pharmacopoeia for endotoxins (ELC), which should be detectable also using the WB/IL-1 test.

Human serum albumin belongs to those substances that still are tested in the rabbit pyrogen test. Spreitzer et al (2002) compared the sensitivity of the rabbit test with the WB/IL-1 using 29 defined human albumin samples: plain, spiked with 5 EU/ml and 10 EU/ml respectively. The unspiked samples were negative in both assays. Both the borderline 5 EU/kg and the 10 EU/kg partially led to results with the rabbit test (conducted with 3 rabbits), which would cause further testing with additional animals. In contrast, the WB/IL-1 test detected 100% detection of the 5 EU/ml and 10 EU/ml endotoxin spikes. The WB/IL-1 demonstrated at least the same level of safety for the products as achieved with the rabbit pyrogen test. After further dilution of the 29 spiked albumin samples to endotoxin levels of 0.5 EU/ ml, 18 samples were still positive in the WB/IL-1 assay but there were 11 negative results too.

Schindler et al. (2003) directly compared the reactivity of human and rabbit blood *in vitro* towards Gram negative and Gram-positive stimuli using an *in vitro* whole blood test

(endpoint; IL-1) for both species. The reactivity of the two species towards LPS was found to be similar, whereas human blood was more sensitive for LTA (lipoteichoic acid) than rabbit blood. The results suggested that the test with human blood to detect contaminations in e.g. parenteral drugs, might predict the human reaction to real life contamination better than the rabbit pyrogen test.
A Gram-positive standard derived form *B. subtilis* has been developed by the same research group (University of Konstanz) and was reviewed in numerous different articles. This lipoteichoic acid is BET negative which however reacts positive in the WB/IL-1 assay. Identification, isolation and purification of other Gram-positive stimuli are subject of ongoing research.
It is stressed throughout these studies using whole blood that only healthy donors not taking any medication must be used for testing pharmaceuticals. Various drugs such as cortisone suppress interleukin release and would therefore exclude a blood donor from the experiment. An infection can stimulate release, as reflected in an increased baseline value or an abnormally high interleukin response. Therefore, the WB/IL-1 test may only be used if samples have first been shown not to cause interference. The blood group of the human donors does not influence the results of the assay.

9.2 Discussion

Comment on and compare the conclusions published in independent peer-reviewed reports or other independent scientific reviews of the proposed test method. The conclusions of such scientific reports and reviews should be compared to the conclusions reached in this submission. Any ongoing evaluations of the proposed test method should be described.
The validation study summarised in this BRD is the first, which extensively addresses specificity and accuracy using actual medicinal products spiked with endotoxin. Hence, there are no comparing reports in independent peer-reviewed journals available. However, the validation study confirms conclusions of several scientific reports, e.g. several preliminary studies (e.g. Jahnke et al.2000, Fennrich et al., 1999, Spreitzer et al 2002) showed that the WB/IL-1 assay is suitable to test different types of pharmaceuticals. Their findings are confirmed by the current validation study, where 11 different pharmaceuticals were tested. In addition, both studies indicate that (pyrogen free) batches which passed the current batch release scheme and are available on the market, show rarely a false positive reactivity in the WB/IL-1 assay. Jahnke's study was conducted by an experienced laboratory, whereas relatively inexperienced laboratories were also involved in the presented validation study. This may account for the less than 100% specificity in the validation study.

Finally, Charles River Endosafe offers the WB/IL-1 test under the name IPT (*In vitro* Pyrogen Test) worldwide in a highly standardized kit-version. Frequent symposia and workshops with coworkers of Charles River together with the University of Konstanz take place in order to train interested parties and introduce the IPT to users. The introduction and optimization of cryopreserved human whole blood is expected to overcome all final obstacles to standardization.

9.3 Results of similar validated method

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the results of studies conducted with the validated test method subsequent to the ICCVAM evaluation should be included and any impact on the reliability and accuracy of the proposed test method should be discussed.

As mentioned, *in vitro* methods activating monocytoid cells for detecting pyrogenic contaminants are being developed over the course of the past two decades. A number of variants have been described, although the underlying principle of each variant remains the same. The test preparation is cultured with monocytoid cells, either as peripheral blood mononuclear cells (PBMCs), (diluted) whole blood or cells of a monocytoid cell line such as MONO MAC-6 (MM6). Accuracy and specificity of these test methods are comparable, but in general methods using whole blood, PBMC and the MM6 cell line appear to perform best (Hoffmann et al, 2005b).

Table 9.3.1 summarises the performance of *in vitro* methods presented in the five BRDs and Table 9.3.2 compares the *in vivo* and *in vitro* pyrogen tests regarding their strengths, weaknesses, costs, time, limitations.

However, most studies (as this one) are done with model pyrogens and as yet little experience is available in the field, e.g. as part of the final batch release test-package. Experience and thus confidence in these methods will grow once regulatory authorities approve these methods and more manufacturers start to employ them. Then, on a case by case situation, it should be determined which method is best suited for the actual situation and demonstrates to pick out the appropriate, i.e. pyrogenic batches of the medicinal product.

Table 9.3.1: Summary of the performance of in vitro pyrogen tests based on monocytoid cells (see Tables 7.2.2; 7.2.4; 6.1.3)

Test	System	Read-out	Intralaboratory reproducibility (%)	Interlaboratory reproducibility (%)	Sensitivity (%)	Specificity (%)
WB/IL-6	whole blood	IL-6	DL: 83.3 NL1: 94.4 NL2: 100	DL-NL1: 85.4 DL-NL2: 85.4 NL1-NL2: 92.0	88.9	96.6
WB/IL-1	whole blood	IL-1β	DL: 88. 9 NL1: 95.8 NL2: 94.4	DL-NL1: 72.9 DL-NL2: 81.6 NL1-NL2: 70.2	72.7	93.2
96-wells WB/IL-1 [1]	whole blood	IL-1β	-	DL-NL1: 88.1 DL-NL2: 89.7 NL1-NL2: 91.5	98.8	83.6
CRYO WB/Il-1	cryo whole blood	IL-1β	-	DL-NL1: 91.7 DL-NL2: 91.7 NL1-NL2: 91.7	97.4	81.4
KN CRYO WB/Il-1 [2]	cryo whole blood	IL-1β	-	DL-NL1: 83.3 DL-NL2: 100 NL1-NL2: 83.3	88.9	94.4
PBMC/IL6	PBMC	IL-6	DL: 94.4 NL1: 100 NL2: 94.4	DL-NL1: 84.0 DL-NL2: 86.0 NL1-NL2: 90.0	92.2	95.0
PBMC-CRYO/IL-6 [3]	PBMC	IL-6	-	DL-NL1: 96 DL-NL2: 76 NL1-NL2: 80	93.3	76.7
MM6/IL-6	MM6	IL-6	DL: 100 NL1: 94.4 NL2: 94.4	DL-NL1: 90.0 DL-NL2: 89.6 NL1-NL2: 83.3	95.5	89.8

DL = developing laboratory; NL1, NL2 = naive laboratory 1 and 2
1 = data provided in Section 13 of WB/IL-1 BRD
2 = data provided in Section 13 of CRYO WB/IL-1 BRD
3 = data provided in Section 13 of PBMC/IL-6 BRD

Table amended from Hoffmann et al 2005b; results with THP cells not included

BRD: CRYO WB/IL-1 March, 2006

Table 9.3.2: Comparison of the in vivo and in vitro pyrogen tests regarding their strengths, weaknesses, costs, time, limitations

	Rabbit pyrogen test	**BET / LAL**	**In vitro pyrogen test**
Test materials	Liquids	Clear liquids	Liquids, potentially cell preparations, solid materials
Pyrogens covered	All (possible species differences to humans for non-endotoxin pyrogens)	Endotoxin from Gram-negative bacteria	(probably) all
Limit of detection (LPS)	0,5 EU	0,1 EU (some variants down to 0,01 EU)	0,5 EU (validated PM), some variants down to 0,001 EU
Ethical concerns	Animal experiment	About 10% lethality to bled horseshoe crabs	Some assays: blood donation
Costs*	High (200-600$/sample)	Low (50-150$/sample)	Medium (100-350$/sample)
Time required	27 h	45 min	24-30h**
Materials not testable	Short-lived radiochemicals, anesthetics, sedatives, analgetics, chemotherapeutics, immunomodulators, cytokines, corticosteroids	Most biologicals, glucan-containing preparations (herbal medicinal products, cellulose-filtered products), lipids, microsomes, cellular therapeutics	Not known (some of the materials not testable in rabbits require adaptations)
Others	No positive or negative control included, strain differences, stress affects body temperature	Potency of LPS from different bacterial species in mammals not reflected, false-positive for glucans	Possible donor differences, need to exclude hepatitis/HIV and acute infections / allergies of donors, dedifferentiation of cell lines

* = We consulted the laboratories participating in the validation study and a consultant regarding the costs of the tests. The figures we received vary significantly depending on the facility (e.g. industry, contract laboratory, control authority), frequency of testing, specific test requirements, country, etc.

** = interference testing might increase duration by 24 hours

10 Animal Welfare Considerations (Refinement, Reduction, and Replacement)

10.1 Diminish animal use

Describe how the proposed test method will refine (reduce or eliminate pain or distress), reduce, or replace animal use compared to the reference test method.

Depending on the medicinal product, one of two animal-based pyrogen tests is currently prescribed by the respective Pharmacopoeias, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET). The rabbit pyrogen test detects various pyrogens but alone the fact that large numbers of animals are required to identify a few batches of pyrogen-containing samples argues against its use. In the past two decades, the declared intention to refine, reduce and replace animal testing, has lowered rabbit pyrogen testing by 80% by allowing to use the BET as an alternative pyrogen test for certain medicinal products. Bacterial endotoxin is the pyrogen of major concern to the pharmaceutical industry due to its ubiquitous sources, its stability and its high pyrogenicity. With the BET, endotoxin is detected by its capacity to coagulate the amoebocyte lysate from the haemolymph of the American horseshoe crab, *Limulus polyphemus*, a principle recognised some 40 years ago (Levin and Bang, 1964). In the US, *Limulus* crabs are generally released into nature after drawing about 20% of their blood and therefore most of these animals survive. However, the procedure still causes mortality of about 30,000 horseshoe crabs per year, which adds to more efficient threats of the horseshoe crab population like its use as bait for fisheries, habitat loss and pollution.

The proposed test method is an alternative for the rabbit test and the BET. By replacing the rabbit test or the BET, the lives of rabbits and horseshoe crabs are spared.

10.2 Continuation of animal use

If the proposed test method requires the use of animals, the following items should be addressed:

10.2.1 *Describe the rationale for the need to use animals and describe why the information provided by the proposed test method requires the use of animals (i.e., cannot be obtained using non-animal methods).*

Not applicable.

10.2.2 *Include a description of the sources used to determine the availability of alternative test methods that might further refine, reduce, or replace animal use for this testing. This should, at a minimum, include the databases searched, the search strategy used, the search date(s), a discussion of the results of the search, and the rationale for not incorporating available alternative methods.*

Not applicable.

10.2.3 *Describe the basis for determining that the number of animals used is appropriate.*

Not applicable.

BRD: CRYO WB/IL-1 March, 2006

10.2.4 *If the proposed test method involves potential animal pain and distress, discuss the methods and approaches that have been incorporated to minimize and, whenever possible, eliminate the occurrence of such pain and distress.*
Not applicable.

11 Practical Considerations

11.1 Transferability

Discuss the following aspects of proposed test method transferability. Include an explanation of how this compares to the transferability of the in vivo reference test method and, if applicable, to a comparable validated test method with established performance standards.

In general, the proposed test method is not unlike other bioassays and immunoassays that are performed routinely in many laboratories.

11.1.1 *Discuss the facilities and major fixed equipment needed to conduct a study using the proposed test method*.

No extraordinary facilities are required. General laboratory equipment for aseptic operations and analytical instruments for performing immunoassays, e.g. microtiter plate reader and –washer, are sufficient to perform the proposed test method.

11.1.2 *Discuss the general availability of other necessary equipment and supplies*.

All supplies and reagents are readily available on the market. In contrast, availability of sufficient rabbits of adequate weight and in good health for the *in vivo* reference test is sometimes reported a limitation.

11.2 Training

Discuss the following aspects of proposed test method training. Include an explanation of how this compares to the level of training required to conduct the in vivo reference test method and, if applicable, a comparable validated test method with established performance standards.

11.2.1 *Discuss the required level of training and expertise needed for personnel to conduct the proposed test method.*

The proposed test method requires personnel trained for general laboratory activities in cell biology and immunochemistry or biochemistry. Techniques they should master are not unlike cell culture (aseptic operations) and immunological techniques (especially ELISA). Such expertise is available in most if not all QC-laboratories.

11.2.2 *Indicate any training requirements needed for personnel to demonstrate proficiency and describe any laboratory proficiency criteria that should be met.*

Personnel should demonstrate that they master the execution of the test. The candidate should demonstrate to meet all the appropriate assay acceptance criteria and yield accurate results (outcome) using selected test items.

11.3 Cost Considerations

Discuss the cost involved in conducting a study with the proposed test method. Discuss how this compares to the cost of the in vivo reference test method and, if applicable, with that of a comparable validated test method with established performance standards.

Three factors contribute to the cost of the proposed test method: availability of monocytoid cells, cost of the reagents for the immunoassay and, last but not least, personnel.

Since the proposed test method is relatively more labor-intensive, it is estimated that the cost of the proposed test method is more then the BET or the *in vivo* reference test using rabbits. Obviously, a higher throughput of tests (runs/year) such as in a QC-laboratory of a multi-product facility or in a Contract Research Organization will significantly reduce the costs per assay.

However, especially with pharmaceuticals of biological origin, the proposed test method may be cost-effective, since these products all to often are incompatible with the BET and by their nature preclude the re-use of the rabbits.

11.4 Time Considerations

Indicate the amount of time needed to conduct a study using the proposed test method and discuss how this compares with the in vivo reference test method and, if applicable, with that of a comparable validated test method with established performance standards.

Esssentially the test stretches two working days. On day one the testing materials are prepared and incubated overnight with the monocytoid cells. On the second day the amount of excreted cytokines is determined by immunoassay. The total time from start to result is approximately 24 hours.

It is thus concluded that the proposed test method will take more time when compared to the alternative tests, either the rabbit test or the BET. It should be noted that rabbits are tested prior to their first use by a sham test.

12 References

List all publications referenced in the submission.

References in bold are included as hardcopies in Appendix B

Barnett, V., Lewis, T. (1984). Outliers in statistical data. 2nd edition, Chichester, US: John Wiley & Sons, pp 171-172.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Clopper, C. J., Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26: 404-413

Co Tui & Schrift, M.H. (1942). A tentative test for pyrogen in infusion fluids. Proc. Soc. Exp. Biol. Med. 49: 320-330.

Cooper JF, Levin J, Wagner Jr. HN (1971). Quantitative comparison of *in vitro* and *in vivo* methods for the detection of endotoxin. J Lab Clin Med 78: 138-145.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Fennrich S, Fischer M, Hartung T, Lexa P, Montag-Lessing T, Sonntag H-G, Weigandt M und Wendel A. (1999). Detection of endotoxins and other pyrogens using human whole blood. Dev. Biol. Standards 101:131-139

Fennrich S, Wendel A and Hartung T. (1999). New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 16:146-149

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005a). Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Pertbauer A, Schindler S, et al. (2005b) Validation of novel pyrogen tests based on human monocytoid cell lines. J Immunol Methods, 298(1-2):161-173.

Hothorn, L. A. (1995). Biostatistical analysis of the 'control vs. k treatments' design including a positive control group. In: Testing principls in clinical and preclinical trials, pp 19-26, Stuttgart, Germany: Gustav Fischer Verlag.

Jahnke M, Weigand, Sonntag H-G. (2000). Comparative testing for pyrogens in parenteral drugs using the human whole blood pyrogen test, the rabbit in vivo pyrogen test and the LAL test. European Journal of Parenteral Science 5(2):39-44.

Levin J, Bang FB (1964). A description of cellular coagulation in the Limulus. Bull Johns Hopkins Hosp 155: 337-345.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, November 1989: 17-18.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' *in vitro* test for pyrogens. J Immunol Methods 274: 209-220.

Schindler S, Bristow A, Cartmell T, Hartung T and Fennrich S. (2003). Comparison of the reactivity of human and rabbit blood towards pyrogenic stimuli. ALTEX 20:59-63

Schindler et al (2004) Cryopreservation of human whole blood for pyrogenicity testing. Journal of Immunological Methods 294, 89–100

Snedecor, G. W., Cochran, W. G. (1989). Statistical methods, 8th edition, Ames, US: The Iowa State University Press.

Spreitzer I, Fischer M, Hartzsch K, Lüderitz-Püschel U and Montag T. (2000). Comparative Study of Rabbit Pyrogen Test and Human whole Blood Assay on Human Serum Albumin. ALTEX 19 (Suppl. 1):73-75.

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

13 In vitro Pyrogen Test with Cryopreserved Human Whole Blood according the Konstanz method (KN CRYO WB/IL-1).

13.1 Rationale

Several laboratories have developed their own method to cryopreserve whole blood. In the main part of this BRD, whole blood cryopreserved according to a procedure developed by PEI is used. Another cryopreserved WB/Il-1 method recently published by Schindler et. al (2004) was also investigated separately, while applying the same study plan. This variant is indicated as the KN (University of Konstanz, Germany) CRYO WB/IL-1 method. The optimisation of the procedure to preserve whole blood is in detail described in the manuscript of Schindler (Schindler et al 2004; Appendix B BRD).

13.2 Test Method Protocol Components

The method follows the original standard protocol, with the obvious exemption of the cryopreservation of the freshly drawn whole blood in the presence of 10% (v/v) DMSO prepared according to the Konstanz method (Schindler et. al, 2004). Details of the test procedure are given in the method protocol under point 7C (Appendix A). The blood has to be stored in the vapour phase of liquid nitrogen. After incubation of the blood (20 µl) with the samples of interest in a 96-wells microtiter plate, the plate is frozen at -20 or -80^{o}C until the contents of the wells are completely frozen. Subsequently the plate is thawed at room temperature or in a water bath (maximum 37^{o}C). The released IL-1□ is assessed using the standard IL-1□ ELISA.

13.3 Substances Used for Validation

The same 10 parenteral drugs used to determine sensitivity and specificity (see table 3.3.1.) were used for the catch-up validation. Again, each test item was tested after spiking at its individual MVD, thus came with its own specific set of 5 endotoxin spike solutions: 0.0, 0.25, 0.5, 0.5 and 1.0 EU/ml. The test items were assessed with 5 different endotoxin levels at 3 independent test facilities, yielding a total of 150 data points, biometrically considered to be sufficient for further analysis.
The same three drugs (table 3.3.2) as used for the prevalidation of the CRYO WB/IL-1 method were employed. Each drug was tested at an interference free dilution and spiked with 0.0, 0.0, 0.5 and 1.0 EU/ml. The samples were tested at each of the 3 laboratories. The results were used to provide a preliminary estimate of the interlaboratory reproducibility and accuracy.

13.4 Preliminary estimate of the Test Method Accuracy

In short, three test substances were spiked with WHO-LPS at four levels (0, 0, 0.5 and 1.0 EU/ml respectively), which resulted in 12 samples with a balanced design with regard to positive and negative samples (i.e. samples expected to be pyrogenic and non-pyrogenic respectively. These 12 samples were tested in three laboratories (See figure 13.4.1)

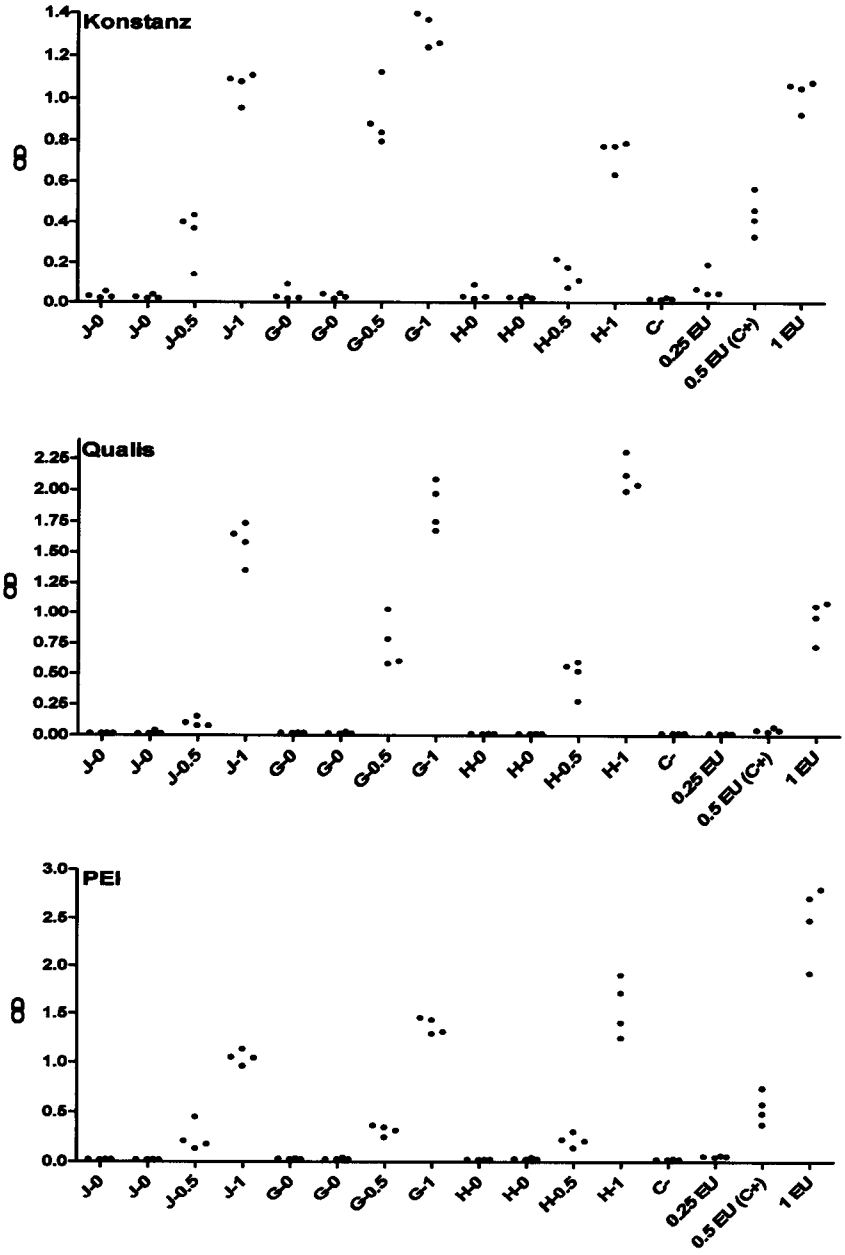


Figure 13.4.1: Prevalidation data for KN CRYO WB/IL-1 of the three involved laboratories. The treatments and controls are abbreviated (indicating the endotoxin contamination in EU. (J = Jonosteril: G = Gelafundin; H = Haemate; C- = saline with 0 EU; C+ = positive control)

As Figure 13.4.1 only gives an indication about variability of replicates, the coefficients of variation (CVs) were calculated for each treatment or control for all laboratories (Figure 13.4.2). While the major part of the CVs was smaller than 40%, six samples (mainly from Konstanz) and one standard showed a CV larger than 45%.

BRD: CRYO WB/IL-1 March, 2006

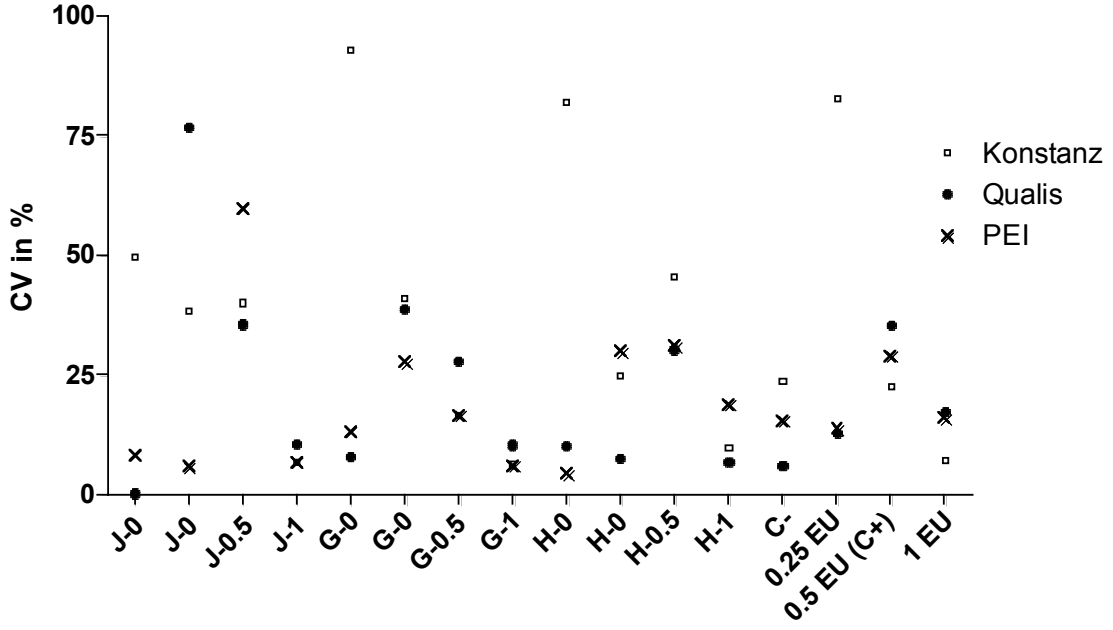


Figure 13.4.2: Coefficients of variation of the prevalidation data from KN WB-CRYO/IL1-B for the three involved laboratories. The treatments and controls are abbreviated indicating the endotoxin contamination in EU. (J = Jonosteril: G = Gelafundin; H = Haemate; C- = saline with 0 EU; C+ = positive control)

Application of the PM to these data resulted in the classifications summarized in Table 13.4.1. Ten out of the twelve spikes were classified in the same way in all laboratories (83.3%). Comparing the laboratories pair wise, showed that 32 of the total of 36 single comparisons, i.e. 88.9% resulted in the same classification.
Assessing in the final step preliminarily the predictive capacity, revealed that one negative samples was classified wrongly (Qualis: J-0) due to one outlying value, and that two times a Haemate 0.5-EU sample (Konstanz and PEI) at the rabbit classification threshold was classified false negative.
These false negative samples had OD-values significantly larger than (PEI) or equivalent to (Konstanz) the respective 0.25 EU-spike of the standard curve.

BRD: CRYO WB/IL-1 March, 2006

Table 13.4.1: Classification by the KN CRYO WB/IL-1 of the spikes in the prevalidation in the three involved laboratories

drug	**spike in EU**	**laboratory Konstanz**	**Qualis**	**PEI**
Jonosteril	0	0	0	0
	0	0	**1**	0
	0.5	1	1	1
	1	1	1	1
Gelafundin	0	0	0	0
	0	0	0	0
	0.5	1	1	1
	1	1	1	1
Haemate	0	0	0	0
	0	0	0	0
	0.5	**0**	1	**0**
	1	1	1	1

Table 13.4.2: Preliminary estimate of interlaboratory reproducibility: Assessed by testing of 3 substances, spiked 4 times. One run of 12 samples by three different laboratories.

Laboratories	**Interlaboratory Reproducibility**	**Number of equal predictions**
DL - NL1	83.3%	10 / 12
DL - NL2	100%	12 / 12
NL1 – NL2	83.3%	10 / 12
Mean	88.9%	
same result in all laboratories	83.3%	10 / 12

DL =Konstanz; NL1 = Qualis; NL2 = PEI

A 2x2 contingency table was constructed (table 13.4.3), from which the estimates of sensitivity and specificity can easily be derived.

Table 13.4.3: 2x2 contingency table. The prediction model applied to a preliminary validation study with KN CRYO WB/IL-1. Three different substances were assessed in three different laboratories (derived from table 13.4.1)

		True status of samples		**Total**
		+	-	
PM	+	16	1	17
	-	2	17	19
	Total	18	18	36

The specifications of specificity and sensitivity described in section 5.3 were applied to these results. The specificity (Sp) of the KN CRYO WB/IL-1 assay is 94.4% (17/(17+1)*100%). The sensitivity (Se) calculated for this data set is 88.9% (16/(16+2) *100%). As outlined previously the specificity is overestimated and the sensitivity is underestimated as a result of the design of this part of the study.

Conclusion: For KN CRYO WB/IL-1, an increased inherent variability and an increased and borderline acceptable limit of detection compared to the WB/IL-1 and the CRYO WB/IL-1 were found. As the variability is still tolerable and because of the interlaboratory reproducibility and the predictive capacity in terms of specificity and sensitivity, the KN CRYO WB/IL-1 method showed sufficiently good results to proceed with the validation of this method.

13.5 Test Method Accuracy

To assess **accuracy** of the proposed test method 10 substances (listed in table 3.3.1), were spiked with five different concentrations of the WHO-LPS (one of which is negative). To permit the application of the chosen PM, each drug was diluted to its individual MVD, which has been calculated using the ELC to that drug (listed in table 3.3.1.) Each substance had their own specific set of endotoxin spike solution, ensuring that after spiking the undiluted drug, it contained 0.0; 0.25; 0.5; 0.5 and 1.0 EU/ml respectively when tested at its MVD. A detailed description of the sample preparation was supplied to the three independent laboratories (as shown in table 5.1.2). To put more weight to the result of this part of the validation, the spikes were blinded and coded by QA ECVAM. Accuracy was assessed by applying the PM to the results and evaluating the concordance in a two by two table.

Table 13.4.1: Results of the validation study of 10 drugs, spiked with WHO-LPS at 0.0, 0.25, 0.5, 0.5 and 1.0 IU/ml, respectively and tested in 3 different laboratories. Samples and spikes were blinded. Classifications after applying the prediction model.

drug (code)	**spike EU/ml**	**"truth"**	**results** PEI	Qualis	Novartis
Beloc (BE)	0.00	0	0	0	0
	0.25	0	0	0	CV
	0.50	1	**0**	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Binotal (BI)	0.00	0	0	0	0
	0.25	0	0	**1**	CV
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Ethanol 13% (ET)	0.00	0	0	0	0
	0.25	0	0	0	CV
	0.50	1	1	1	**0**
	0.50	1	1	1	**0**
	1.00	1	1	1	1
Fenistil (FE)	0.00	0	0	0	0
	0.25	0	0	CV	**1**
	0.50	1	CV	1	CV
	0.50	1	1	1	1
	1.00	1	1	1	1
Glucose 5% (GL)	0.00	0	0	0	0
	0.25	0	0	**1**	0
	0.50	1	**0**	1	CV
	0.50	1	**0**	1	1
	1.00	1	**0**	1	1
"Drug A"	0.00	0	0	0	nq
0.9% NaCl (LO)	0.25	0	0	0	nq
	0.50	1	**0**	1	nq
	0.50	1	**0**	1	nq
	1.00	1	1	1	nq
MCP (ME)	0.00	0	0	0	0
	0.25	0	0	**1**	0
	0.50	1	**0**	1	CV
	0.50	1	**0**	1	1
	1.00	1	1	1	1

BRD: CRYO WB/IL-1 March, 2006

drug (code)	**spike EU/ml**	**"truth"**	**results** PEI	Qualis	Novartis
"Drug B"	0.00	0	0	0	nq
0.9% NaCl (MO)	0.25	0	0	0	nq
	0.50	1	**0**	1	nq
	0.50	1	**0**	1	nq
	1.00	1	1	1	nq
Orasthin (OR)	0.00	0	0	0	nq
Syntocinon	0.25	0	CV	CV	nq
	0.50	1	1	1	nq
	0.50	1	1	1	nq
	1.00	1	1	1	nq
Sostril (SO)	0.00	0	0	0	nq
	0.25	0	0	**1**	nq
	0.50	1	**0**	1	nq
	0.50	1	1	1	nq
	1.00	1	1	1	nq

"0"denotes "non-pyrogenic"; "1" denotes "pyrogenic".
Grey shading indicates that for these drugs the PPCs did not qualify so that the PC was used in the PM.
CV = sample showed a variability resulting in exclusion, i.e. CV > 45% and no significant outlier present.
nq = not qualified according to quality criteria, i.e. failure of PPCs and PCs
False classifications are in bold/color type.

Of the 150 available data for the KN CRYO WB/IL-1 method, eleven sets of 4 replicates showed a high variability resulting in exclusion, i.e. CV > 45% and no significant outliers present. Therefore 139 data in total could be used to estimate the specificity and sensitivity of the KN CRYO WB/IL-1 method. The results are shown separately for each participating laboratory (table 13.5.1) as well as combined for all these laboratories (table 13.5.2).

The specificity (table 13.5.3) that can be estimated from the available results for DL, NL1 and NL2 is 94.1%, 80% and 77.8% respectively The estimated sensitivity (table 13.5.3) of the KN CRYO WB/IL-1 assay was excellent of all three participating laboratories: 96%, 100% en 100% respectively (calculated from results in table 13.5.2).

Table 13.5.1: 2x2 contingency table. Prediction model applied to the KN CRYO WB/IL-1 test result of 10 different substances assessed in three different laboratories. Results of each laboratory separately (DL, NL1 and NL2= PEI, Qualis and Novartis respectively).

Results DL		**True status of samples** +	-	**Total**
PM	+	18	0	18
	-	11	19	30
Total		29	19	48

Results NL1		**True status of samples** +	-	**Total**
PM	+	30	4	34
	-	0	14	14
Total		30	18	48

Results NL2		**True status of samples** +	-	**Total**
PM	+	13	1	14
	-	2	8	10
Total		15	9	24

Table 13.5.2: 2x2 contingency table. Prediction model applied to the KN CRYO WB/IL-1 test result of 10 different substances assessed in three different laboratories (from table 13.4.1).

	True status of samples +	-	**Total**
PM +	61	5	**66**
-	13	41	**54**
Total	**74**	**46**	**120**

The overall specificity of the KN CRYO WB/IL-1 assay is 89.1% (46/(46+14)*100. The overall sensitivity equals 82.4% (61/(61+13) *100%). Within the laboratories, specificity varied from 77.8%, 88.9% and up to 100%, whereas the sensitivity varied from 62.1% up to 86.7% and 100%. respectively

Table 13.5.3: Specificity and sensitivity of the KN WB-CRYO /IL-1 assay as determined from table 13.5.2

	N total	N correctly identified	proportion	95% CI lower limit	95% CI upper limit
Specificity (Sp)	46	41	89.1%	0.764	0.964
Sensitivity (Se)	74	61	82.4%	0.718	0.903

13.6 Test Method Reliability (Reproducibility)

The interlaboratory reproducibility (table 13.6.1) of the KN CRYO WB/IL-1 method was assessed from the results of the validation test with 10 substances spiked with 5 separate spikes. The reproducibility varied from 68.1% to 82.6% between two laboratories. The estimated reproducibility between three laboratories (65.2%) was based on a very limited number of samples, because in one of the laboratories the results of many samples were not qualified for analysis.

Table 13.6.1: Interlaboratory reproducibility: Assessed by testing of 10 substances, spiked 5 times. One run of 50 samples by three different laboratories.

Laboratories	Interlaboratory Reproducibility	Number of equal predictions
DL - NL1	68.1%	32 / 47
DL - NL2	70.8%	17 / 24
NL1 – NL2	82.6%	19 / 23
Mean	73.8%	
same result in all laboratories	65.2%	15 / 23

DL =PEI; NL1 = Qualis; NL2 = Novartis

13.7 Summary and conclusion

In this study the specificity of the KN CRYO WB/IL-1 is comparable with the CRYO WB/IL-1 (89.1% versus 83.6%; see section 6 of this BRD). However, the mean sensitivity shown for the KN CRYO WB/IL-1 (82.4%) is significantly less then achieved with the CRYO WB/IL-1 (98.8%). This can be fully explained by the false negatives in one of the laboratories. The KN CRYO WB/IL-1 method shows a lower interlaboratory reproducibility then was estimated for the CRYO WB/IL-1 method. In addition, it is noted that in one of the laboratories the results of many samples were not qualified for analysis. It appears that in this study the KN CRYO WB/IL-1 is not yet as robust as the CRYO WB/IL-1 method which is described in sections 1-12 of this BRD.

14 Supporting Materials (Appendices)

14.1 Standard operating procedure (SOP) of the proposed method

Provide the complete, detailed protocol for the proposed test method.

Appendix A includes the test method protocol *CRYO WB/IL-1: Human Whole Blood Pyrogen Test in 96-well Plates Using Fresh or Cryopreserved Blood(electronic file name: SOP CRYO WB IL 1*). It covers three variations to the preparation of the whole blood described under point 7: 7A - fresh blood using 96-well plates, 7B -cryopreserved blood according to the so called "PEI (Paul-Ehrlich-Institute, Langen, Germany)" method = CRYO WB/IL-1" and 7C - cryopreserved blood according to the so called "Konstanz" method KN CRYO WB/IL-1.

The trial plan of the catch-up validation study is also included in Appendix A.

14.2 Standard operating Procedure (SOP) of the reference method

Provide the detailed protocol(s) used to generate reference data for this submission and any prtocols used to generate validation data that differ from the proposed protocol.

14.3 Publications

Provide copies of all relevant publications, including those containing data from the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method with established performance standards.

Not all of the publications cited in the BRD (see section 12) are included as hardcopies in Appendix B, e.g. publications on statistical methods are not given.

Remark: The same set of hardcopies was included into Appendix B of all of the 5 submitted BRDs. Therefore some of the publications in Appendix B might not be referenced in the current BRD nor included in the list of the publications in section 12. Three general publications were added, which are not cited in any BRD but might be useful as background information to the validation study: the ECVAM publications on validation (Balls et al, 1995; Curren et al, 1995; Hartung et al, 2004). Several publications were included, which either give more background information on the human fever reaction or report on specific studies using *in vitro* pyrogen tests.

List of hard copies

Andrade SS, Silveira RL, Schmidt CA, Junior LB, Dalmora SL. (2003) Comparative evaluation of the human whole blood and human peripheral blood monocyte tests for pyrogens. Int J Pharm. Oct 20;265(1-2):115-24.

Balls et al. (1995) Practical aspects of the validation of toxicity test procedures. ECVAM Workshop Report 5. ATLA 23, 129-147.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Carlin G, Viitanen E. (2003) In vitro pyrogenicity of a multivalent vaccine: Infanrix. PHARMEUROPA, 15(3), 418-423.

Curren et al (1995) The Role of prevalidation in the development, validation and acceptance of Alternative Methods. ECVAM Prevalidation Task Force Report 1. ATLA 23, 211-217.

De Groote, D., Zangerle, P.F., Gevaert, Y., Fassotte, M.F., Beguin, Y., Noizat-Pirenne, F., Pirenne, J., Gathy, R., Lopez, M., Dehart I., (1992). Direct stimulation of cytokines (IL-1 beta, TNF-alpha, IL-6, IL-2, IFN-gamma and GM-CSF) in whole blood. I. Comparison with isolated PBMC stimulation. Cytokine 4, 239.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Fennrich S, Wendel A and Hartung T. (1999). New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 16:146-149

Fennrich S, Fischer M, Hartung T, Lexa P, Montag-Lessing T, Sonntag H-G, Weigandt M und Wendel A. (1999). Detection of endotoxins and other pyrogens using human whole blood. Dev. Biol. Standards 101:131-139

Gaines Das et al (2004). Monocyte activation test for pro-inflammatory and pyrogenic contaminants of parenteral drugs: test design and data analysis. J. of Immunological Methods 288, 165-177.

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hartung et al (2004) A modular approach to the ECVAM principles on test validity. ATLA 32, 467-472.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005). Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Peterbauer A, Schindler S, et al (2005). International validation of novel pyrogen tests based on human monocytoid cells. J Immunol Methods 298: 161-173.

Jahnke M, Weigand, Sonntag H-G. (2000). Comparative testing for pyrogens in parenteral drugs using the human whole blood pyrogen test, the rabbit in vivo pyrogen test and the LAL test. European Journal of Parenteral Science 5(2):39-44

Nakagawa Y, Maeda H, Murai T. (2002) Evaluation of the in vitro pyrogen test system based on proinflammatory cytokine release from human monocytes: comparison with a human whole blood culture test system and with the rabbit pyrogen test. Clin Diagn Lab Immunol. May;9(3):588-97.

Pasare, C. and Medzhitov, R. (2003). Toll pathway-dependent blockade of CD4+CD25+ T cell-mediated suppression by dendritic cells. Science 299: 1033-1036.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1998) The detection of pyrogens in blood products using an ex vivo whole blood culture assay. J Immunoassay. May-Aug;19(2-3):95-111.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1999) Differentiation between endotoxin and non-endotoxin pyrogens in human albumin solutions using an ex vivo whole blood culture assay. J Immunoassay. Feb-May; 20(1-2):79-89.

Poole S, Thorpe R, Meager A, Hubbard AJ, Gearing AJ (1988). Detection of pyrogen by cytokine release. Lancet 8577, 130.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, November 1989: 17-18.

Poole S, Musset MV (1989). The International Standard for Endotoxin: evaluation in an international collaborative study. J Biol Stand 17: 161-171.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' in vitro test for pyrogens. J Immunol Methods 274: 209-220.

Ray A, Redhead K, Selkirk S, Polle S (1991) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211-218.

Reddi, K., Nair, S.P. et al (1996). Surface-associated material from the bacterium Actinobacillus actinomycetemcomitans contains a peptide which, in contrast to lipopolysaccharide, directly stimulates fibroblast interleukin-6 gene transcription. Eur. J. Biochem. 236: 871-876.

Schindler S, Bristow A, Cartmell T, Hartung T and Fennrich S. (2003). Comparison of the reactivity of human and rabbit blood towards pyrogenic stimuli. ALTEX 20:59-63.

Schindler et al (2004) Cryopreservation of human whole blood for pyrogenicity testing. Journal of Immunological Methods 294, 89–100

Schins RP, van Hartingsveldt B, Borm PJ. Ex vivo cytokine release from whole blood. A routine method for health effect screening. Exp Toxicol Pathol. 1996 Nov; 48(6):494-6.

Spreitzer I, Fischer M, hartzsch K, Lüderitz-Püschel U and Montag T. (2000). Comparative Study of Rabbit Pyrogen Test and Human whole Blood Assay on Human Serum Albumin. ALTEX 19 (Suppl. 1):73-75

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

Part 2:

List of Diploma theses, reports and/or PhDs etc. concerning the WB/IL-1 test (IPT: In vitro Pyrogen Test)

1. Final report for the BMBF (Bundesministerium für Bildung und Forschung) (University of Konstanz, 2000). „Evaluierung und Prävalidierung eines Vollblutmodelles zum Ersatz des Pyrogentests am Kaninchen (DAB10)“, Phase I, (“Evaluation and prevalidation of a whole blood assay for the replacement of the pyrogentest with rabbits”), July 1th, 1997 – June 30th, 2000, No. 0311424
2. Final report for the BMBF (Bundesministerium für Bildung und Forschung) (Langen, Paul-Ehrlich-Institut, 2000). „Evaluierung und Prävalidierung eines Vollblutmodelles zum Ersatz des Pyrogentests am Kaninchen (DAB10)“, Phase I, (“Evaluation and prevalidation of a whole blood assay for the replacement of the pyrogentest with rabbits”), July 1th, 1997 – June 30th, 2000, No. 0311425
3. PhD-Thesis from Markus Weigandt at the Ruprecht-Karls-University of Heidelberg, institute of hygiene (Director: Prof. H.-G. Sonntag): Der humane Vollblut-Pyrogentest: Optimierung, Validierung und Vergleich mit den Arzneibuchmethoden” (The human whole blood pyrogen test: optimization, validation and comparision with methods regulated in the pharmacopoeias), 2000
4. Master Thesis (Master of Science: MSc), Karin Kullmann: „Adaptation des *In vitro* Pyrogen Tests (IPT) für prothetische Materialien“ (“Adaptation of the *in vitro* pyrogen test (IPT) to medical devices”), Technical University of Furtwangen, July 2002
5. Final report for the BMBF (Bundesministerium für Bildung und Forschung) (Langen, Paul-Ehrlich-Institut, 2004). „Ersatz des Pyrogentests am Kaninchen durch einen Vollbluttest“, Phase II, (“replacement of the rabbit experiment with the whole blood test”), October 1th, 2000 – September 30th, 2003, No. 0311424A
6. Final report for the BMBF (Bundesministerium für Bildung und Forschung) (University of Konstanz, 2004). „Ersatz des Pyrogentests am Kaninchen durch einen Vollbluttest“, Phase II, (“replacement of the rabbit experiment with the whole blood test”), September 1th, 2000 – August 31th, 2003, No. 0311424A
7. Brazil/Germany Cooperation Project: final report for the BMBF (Bundesministerium für Bildung und Forschung). „Validation of *in vitro* Cytokine Release Assay (Whole Blood Assay) for Controlling the Quality of Human Injectable Products“ for bilaterial Cooperation in Science and Technology (Germany – Brazil), April 1th 2002-March 31th 2004, No. BRA 02/004
8. Cuba/Germany Cooperation Project: final report for the BMBF (Bundesministerium für Bildung und Forschung). „Pyrogenicity Testing by Human Whole Blood“ for bilaterial Cooperation in Science and Technology (Germany – Cuba), January 1th , 2001- December 31th, 2003, No. CUB 00/022
9. Final report for the BMWa (Bundesministerium für Wirtschaft und Arbeit): “Entwicklung einer humanrelevanten Messtechnik für luftgetragene Toxine mit humanem Vollblut“ (development of a human relevant measurement for air-borne

toxins with human whole blood), Sept 3th 2001– Sept. 30th 2003, No. KF 0317101KRF1

10. Postdoctoral lecture qualification (Habilitation), Bert Zucker, "Luftgetragene Endotoxine in Tierställen" ("air-borne pyrogens in a stable"), Institut für Tier- und Umwelthygiene an der freien Universität Berlin, Berlin, 2004
11. Manuscript for the DIF (Deutsches Industrieforum, DIF-Fachtagung), Stefan Fennrich: "Pyrogenverunreinigungen an medizinischen Oberflächen. *In vitro* pyrogen-Test (IPT) als humanrelevantes Prüfverfahren" (Contamination with pyrogens on medical surfaces: the *in vitro* pyrogen test (IPT) as a human specific method), Würzburg, June 21th -22 th, 2004, No. DIF 21/78/FE

Part 3: Further publications concerning the WB/IL-1 test (IPT)

1. Hartung T und Wendel A. Die Erfassung von Pyrogenen in einem humanen Vollblutmodell. ALTEX 1995,12:70-75
2. Fennrich S, Fischer M, Hartung T, Lexa P, Montag-Lessing T, Sonntag H-G, Weigandt M und Wendel A. Entwicklung und Evaluierung eines Pyrogentests mit menschlichem Blut. ALTEX 1998, 15:123-128
3. Fennrich S, Berthold S, Weigandt M, Lexa P, Sonntag H-G, Hartung T, Wendel A. Tagungsberichte, Pyrogentestung mit humanem Blut. Der Tierschutzbeauftragte 2, 1999, 102-107
4. Bonenberger J, Diekmann W, Fennrich S, Fischer M, Friedrich A, Hansper M, Hartung T, Jahnke M, Löwer J, Montag T, Petri E, Sonntag H-G, Weigand M, Wendel A, Zucker B. Pyrogentestung mit Vollblut. Zusammenfassung eines Status Workshops am Paul-Ehrlich-Institut, Langen, am 22.11.1999. Springer Verlag, Bundesgesundheitsbl-Gesundheitsforsch-Gesundheitsschutz, 2000, 43:525-533
5. Petri E, van de Ploeg A, Habermaier B und Fennrich S. Improved detection of pyrogenic substances on polymer surfaces with an ex vivo human whole-blood assay in comparison to the Limulus amoebocyte lysate test. In: Progress in the Reduction, Refinement and Replacement of Animal Experimentation. Editors: Balls M, van Zeller A-M, Halder M.E., Elsevier Science, 2000, 339-345
6. Hartung T, Fennrich S, Fischer M, Montag-Lessing T und Wendel A. Prevalidation of an Alternative to the rabbit test based on human whole blood. In: Progress in the Reduction, Refinement and Replacement of Animal Experimentation. Editors: Balls M, van Zeller A-M, Halder M.E., Elsevier Science, 2000, 991-999
7. Fennrich S, Zucker Bert and Hartung T. Beispiel eines neuen Einsatzbereichs des humanen Vollbluttests: Entwicklung eines Messverfahrens zur Abschätzung der gesundheitlichen Gefährdung durch luftgetragene mikrobielle Verunreinigungen. ALTEX 2001, 18:41-46
8. Thomas Hartung, Ingeborg Aaberge, Susanne Berthold, Gunnar Carlin, Emmanuelle Charton, Sandra Coecke, Stefan Fennrich, Matthias Fischer, Martin Gommer, Marlies Halder, Kaare Haslov, Michael Jahnke, Thomas Montag-Lessing, Stephen Poole, Leonard Schechtman, Albrecht Wendel and Gabriele Werner-Felmayer. Novel

Pyrogen Tests Based on the Human Fever Reaction, The report and Recommendations of ECVAM Workshop 43, 2001, ATLA 29, 99-123

9. Fennrich S, Atemluft, gesund oder gefährlich…..das ist hier die Frage! Tagungsberichte. ALTEX 2002, 19: 43-45
10. Hartung T. Comparison and validation of novel pyrogen tests based on the human fever reaction. ATLA 2002, 30 (Suppl. 2):49-51
11. Morath S, Stadelmaier A, Geyer A, Schmidt RR and Hartung T. Synthetic lipoteichoic acid from Staphylococcus aureus is a potent stimulus of cytokine release. J. Exp. Med., 2002, 195:1635-1640
12. Morath S, Geyer A, Spreitzer I, Hermann C and Hartung T. Structural decomposition and heterogeneity of commercial lipoteichoic acid preparation. Infect. Immun. 2002, 70:938-944
13. Kindinger I, Fennrich S, Zucker B, Linsel G and Hartung T. Determination of air-borne pyrogens by the *in vitro* pyrogen test (IPT) based on human whole blood cytokine response. VDI-Bericht 1656 2002, 499-507
14. Schindler S, Reichstein S, Kindinger I, Hartung T, Fennrich S. New Ways in Pyrogen Testing: Replacing the Rabbit Experiment. Screening, Trends in Drug Discovery May, GIT Verlag, 2-3/2003, 4: 51-53
15. Zucker B A, Linsel G, Fennrich S, Müller W. Die Charakterisierung der entzündungsauslösenden Potenz von Bioaerosolen mittels Interleukinfreisetzung aus humanem Vollblut. Springer, VDI-Verlag. Gefahrstoffe Reinhaltung der Luft (Air Quality Control) 4, 2004, 155-158

14.4 Original data

Include all available non-transformed original data for both the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method with established performance standards.

NOTE: The original data of the ELISA-plate reader were collected by S.Hoffman and ECVAM. These are available on the CD which goes with the BRD.

14.5 Performance standards

If appropriate performance standards for the proposed test method do not exist, performance standards for consideration by NICEATM and ICCVAM may be proposed. Examples of established performance standards can be located on the ICCVAM / NICEATM web site at http://iccvam.niehs.nih.gov.

BRD: CRYO WB/IL-1 March, 2006

APPENDIX A

Trial plan "Catch-up Validation of Novel Pyrogen Tests Based on the Human Fever Reaction"

Detailed protocol CRYO WB/IL-1: *Human Whole Blood Pyrogen Test in 96-well Plates Using Fresh or Cryopreserved Blood(electronic file name: SOP CRYO WB IL 1*).

BRD: CRYO WB/IL-1 March, 2006

APPENDIX B

Not all of the publications cited in the BRD (see section 12) are included as hardcopies in Appendix B, e.g. publications on statistical methods are not given.

Remark: The same set of hardcopies was included into Appendix B of all of the 5 submitted BRDs. Therefore some of the publications in Appendix B might not be referenced in the current BRD nor included in the list of the publications in section 12. Three general publications were added, which are not cited in any BRD but might be useful as background information to the validation study: the ECVAM publications on validation (Balls et al, 1995; Curren et al, 1995; Hartung et al, 2004). Several publications were included, which either give more background information on the human fever reaction or report on specific studies using *in vitro* pyrogen tests.

List of hard copies

Andrade SS, Silveira RL, Schmidt CA, Junior LB, Dalmora SL. (2003) Comparative evaluation of the human whole blood and human peripheral blood monocyte tests for pyrogens. Int J Pharm. Oct 20;265(1-2):115-24.

Balls et al. (1995) Practical aspects of the validation of toxicity test procedures. ECVAM Workshop Report 5. ATLA 23, 129-147.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Carlin G, Viitanen E. (2003) In vitro pyrogenicity of a multivalent vaccine: Infanrix. PHARMEUROPA, 15(3), 418-423.

Curren et al (1995) The Role of prevalidation in the development, validation and acceptance of Alternative Methods. ECVAM Prevalidation Task Force Report 1. ATLA 23, 211-217.

De Groote, D., Zangerle, P.F., Gevaert, Y., Fassotte, M.F., Beguin, Y., Noizat-Pirenne, F., Pirenne, J., Gathy, R., Lopez, M., Dehart I., (1992). Direct stimulation of cytokines (IL-1 beta, TNF-alpha, IL-6, IL-2, IFN-gamma and GM-CSF) in whole blood. I. Comparison with isolated PBMC stimulation. Cytokine 4, 239.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Fennrich S, Wendel A and Hartung T. (1999). New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 16:146-149

Fennrich S, Fischer M, Hartung T, Lexa P, Montag-Lessing T, Sonntag H-G, Weigandt M und Wendel A. (1999). Detection of endotoxins and other pyrogens using human whole blood. Dev. Biol. Standards 101:131-139

Gaines Das et al (2004). Monocyte activation test for pro-inflammatory and pyrogenic contaminants of parenteral drugs: test design and data analysis. J. of Immunological Methods 288, 165-177.

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hartung et al (2004) A modular approach to the ECVAM principles on test validity. ATLA 32, 467-472.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005). Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Peterbauer A, Schindler S, et al (2005). International validation of novel pyrogen tests based on human monocytoid cells. J Immunol Methods 298: 161-173.

Jahnke M, Weigand, Sonntag H-G. (2000). Comparative testing for pyrogens in parenteral drugs using the human whole blood pyrogen test, the rabbit in vivo pyrogen test and the LAL test. European Journal of Parenteral Science 5(2):39-44

Nakagawa Y, Maeda H, Murai T. (2002) Evaluation of the in vitro pyrogen test system based on proinflammatory cytokine release from human monocytes: comparison with a human whole blood culture test system and with the rabbit pyrogen test. Clin Diagn Lab Immunol. May;9(3):588-97.

Pasare, C. and Medzhitov, R. (2003). Toll pathway-dependent blockade of CD4+CD25+ T cell-mediated suppression by dendritic cells. Science 299: 1033-1036.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1998) The detection of pyrogens in blood products using an ex vivo whole blood culture assay. J Immunoassay. May-Aug;19(2-3):95-111.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1999) Differentiation between endotoxin and non-endotoxin pyrogens in human albumin solutions using an ex vivo whole blood culture assay. J Immunoassay. Feb-May; 20(1-2):79-89.

Poole S, Thorpe R, Meager A, Hubbard AJ, Gearing AJ (1988). Detection of pyrogen by cytokine release. Lancet 8577, 130.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of

the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, November 1989: 17-18.

Poole S, Musset MV (1989). The International Standard for Endotoxin: evaluation in an international collaborative study. J Biol Stand 17: 161-171.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' in vitro test for pyrogens. J Immunol Methods 274: 209-220.

Ray A, Redhead K, Selkirk S, Polle S (1991) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211-218.

Reddi, K., Nair, S.P. et al (1996). Surface-associated material from the bacterium Actinobacillus actinomycetemcomitans contains a peptide which, in contrast to lipopolysaccharide, directly stimulates fibroblast interleukin-6 gene transcription. Eur. J. Biochem. 236: 871-876.

Schindler S, Bristow A, Cartmell T, Hartung T and Fennrich S. (2003). Comparison of the reactivity of human and rabbit blood towards pyrogenic stimuli. ALTEX 20:59-63.

Schindler et al (2004) Cryopreservation of human whole blood for pyrogenicity testing. Journal of Immunological Methods 294, 89–100

Schins RP, van Hartingsveldt B, Borm PJ. Ex vivo cytokine release from whole blood. A routine method for health effect screening. Exp Toxicol Pathol. 1996 Nov; 48(6):494-6.

Spreitzer I, Fischer M, hartzsch K, Lüderitz-Püschel U and Montag T. (2000). Comparative Study of Rabbit Pyrogen Test and Human whole Blood Assay on Human Serum Albumin. ALTEX 19 (Suppl. 1):73-75

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

BRD: CRYO WB/IL-1 March, 2006

APPENDIX C

List of abbreviations and definitions

Accuracy	The ability of a test system to provide a test result close to the accepted reference value for a defined property.
BET	The bacterial endotoxin test is used to detect or quantify endotoxins of gram-negative bacterial origin using amoebycte lysate from horseshoe crab (*Limulus polyphemus* or *Tachypleus tridentatus*
BRD	Background Review Document
CRYO WB/IL-1	Whole blood assay (using cryopreserved blood) with IL-1 as endpoint
CV	coefficient of variation
DL	Developing laboratory = laboratory which developed the method or the most experienced laboratory
ELC	Endotoxin limit concentration; maximum quantity of endotoxin allowed in given parenterals according to European Pharmacopoeia
Endotoxins	Endotoxins are a group of chemically similar cell-wall structures of Gram-negative bacteria, i.e. lipopolysaccharides
ELISA	Enzyme linked immunosorbent assay
EU/ml	European Units per ml
IL-1	interleukin 1
IL-6	interleukin 6
Intralaboratory reproducibility	A determination of the extent that qualified people within the same laboratory can independently and successfully replicate results using a specific protocol at different times.
Interlaboratory reproducibility	A measure of the extent to which different qualified laboratories, using the same protocol and testing the same substances, can produce qualitatively and quantitatively similar results. Inter-laboratory reproducibility is also referred to as between-laboratory reproducibility.
KN	University of Konstanz (Konstanz, Germany), developing laboratory WB/IL-1 and CRYO WB/IL-1
LPS	lipopolysaccharides
MM6	MONO MAC-6 cell line

MM6/IL-6	In vitro pyrogen test using MM6 cell line and IL-6 release as an endpoint
MVD	Maximum valid dilution; the MVD is the quotient of the ELC and the detection limit
NIBSC	National Institute for Biological Standards and Control (London, UK), developing laboratory for WB/IL-6
NL	naïve laboratory = laboratory with non or minor experience with the method
NPC	negative product control (clean, released lot of the nominated product under test)
Novartis	Novartis (Basel, Switzerland), developing laboratory PBMC/IL-6
OD	optical density
PBMC	Peripheral blood mononuclear cells
PBMC/IL-6	In vitro pyrogen test using fresh peripheral blood mononuclear cells and IL-6 release as endpoint
PBMC-CRYO/IL-6	In vitro pyrogen test using cryopreserved peripheral blood mononuclear cells and IL-6 release as endpoint
PEI	Paul-Ehrlich Institut (Langen, Germany), participating laboratory
PM	prediction model = is an explicit decision-making rule for converting the results of the in vitro method into a prediction of in vivo hazard
PPC	positive product control (product under test spiked with 0.5 EU/ml of WHO-LPS (code 94/580)
Prevalidation study	A prevalidation study is a small-scale inter-laboratory study, carried out to ensure that the protocol of a test method is sufficiently optimised and standardised for inclusion in a formal validation study. According to the ECVAM principles, the prevalidation study is divided into three phases: protocol refinement, protocol transfer and protocol performance (Curren et al, ATLA 23, 211-217).
Pyrogens	fever-causing materials
Pyrogens, endogenous	endogenous pyrogens are messenger substances released by blood cells reacting to pyrogenic materials; e.g. IL-1, IL-6, TNF-α, prostaglandin E_2
Pyrogens, exogenous	exogenous pyrogens derive from bacteria, viruses, fungi or from the host himself
Reliability	Measures of the extent to which a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is

	assessed by calculating intra- and interlaboratory reproducibility and intra-laboratory repeatability.
Relevance	Relevance of a test method describes whether it is meaningful and useful for a particular purpose. It is the extent to which the measurement result and uncertainty can accurately be interpreted as reflecting or predicting the biological effect of interest.
Repeatibility	Repeatability describes the closeness of agreement between test results obtained within a single laboratory when the procedure is performed independently under repeatability conditions, i.e. in a set of conditions including the same measurement procedure, same operator, same measuring system, same operating conditions and same location, and replicated measurements over a short period of time.
RIVM	National Institute of Public Health and the Environment (Bilthoven, The Netherlands), developing laboratory MM6/IL-6 method
Sensitivity	Sensitivity is the proportion of all positive/active substances that are correctly classified by a test method.
Specificity	Specificity is proportion of all negative/inactive substances that are correctly classified by a test method.
TMB	chromogenic substrate 3,3´,5,5´ -tetramethylbenzidine
TNF-α	tumour necrosis factor-α
USP	US Pharmacopoeia
Validation	Validation is the process by which the reliability and relevance of a procedure are established for a specific purpose
Validation study	A validation study is a large-scale interlaboratory study, designed to assess the reliability and relevance of an optimised method for a particular purpose
WB/IL-1	Whole blood assay (using fresh blood) with IL-1 release as endpoint
WB/IL-6	Whole blood assay (using fresh blood) with IL-6 release as endpoint
WHO	World Health Organization

BRD: CRYO WB/IL-1 March, 2006

Appendix D

Experiments performed regarding the assessment of the intralaboratory reproducibility of CRYO WB/IL-1, KN CRYO WB/IL-1 and 96-wells WB/IL-1

The intralaboratory reproducibility of WB/IL-1 was extensively assessed and the studies carried out for this purpose are described in Section 5 and table 5.1.1 of the BRD.

Table 5.1.1 of WB/IL-1 BRD

Experiment	spikes	n (per spike)	repetitions	
1A	0; 0.5	32	1	64
1B	0; 0.063; 0.125; 0.25; 0.5	12	1	60
2A	0; 0.5	12	3	72
2B	0; 0.25; 0.5	8	3	72
2C	0; 0.5	5	8	80

1A: replicates of the negative and the positive control
1B: limit of detection
2A: behaviour of one fresh blood donor on three successive days
2B: Influence of the operator
2C: Robustness with regard to different donors

For the catch-up validation study, only a few of the experiments described in Table 5.1.1 were carried out (1A and 1B) with some of the methods. The results are given in the following:

1. Comparison of the WB/IL-1 and 96-wells WB/IL-1 (Fig. 1a/b).

Numbers: coefficient of variation (%) as determined using GraphPadPrism Software

BRD: CRYO WB/IL-1 March, 2006

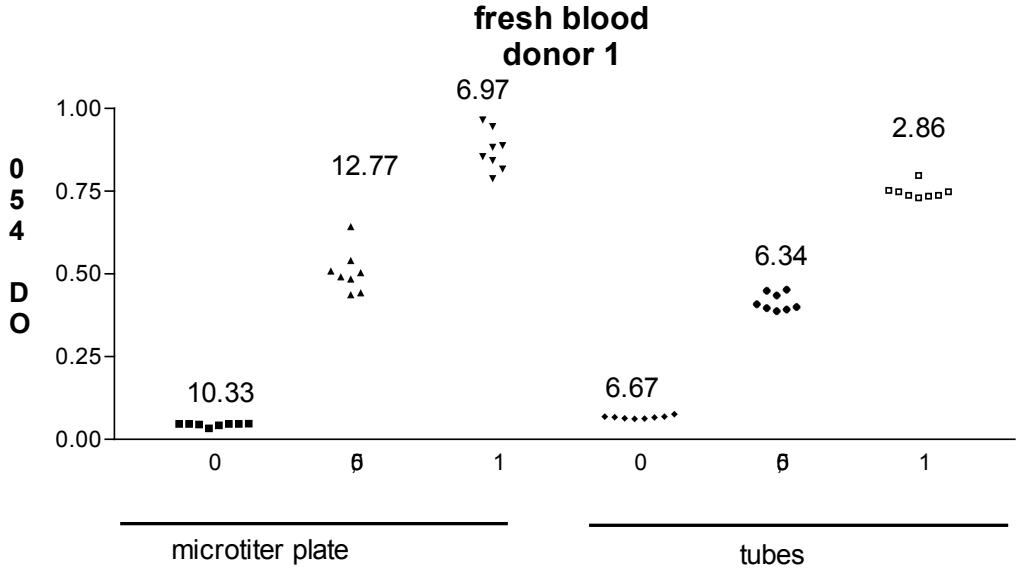


Fig. 1a: Comparison of 96-wells WB/IL-1 with fresh blood of two donors with WB/IL-1 (donor 1)

Fig. 1b: Comparison of 96-wells WB/IL-1 with fresh blood of two donors with WB/IL-1 (donor 2)

BRD: CRYO WB/IL-1 March, 2006

2. With several lots of cryopreserved blood, experiments concerning the variability of results and the detection of the required 0.5 EU/ml stimulus were made (8fold values) and are given in the following figures:

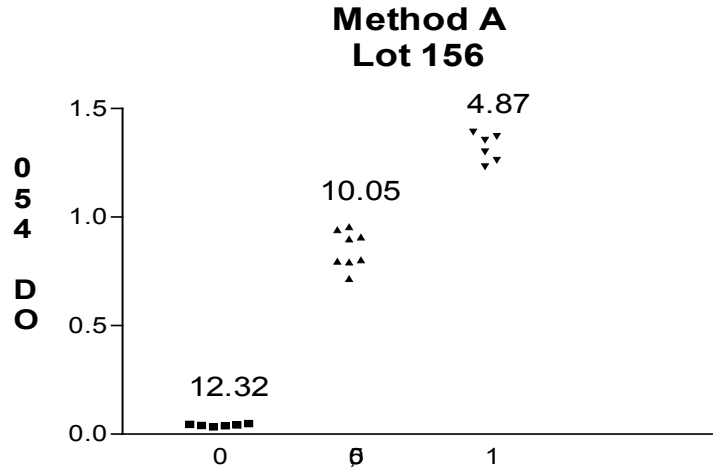


Fig 2a: Method A = CRYO WB/IL-1: blood frozen at -80°C;
Lot 156 8fold tested, 3 spikes

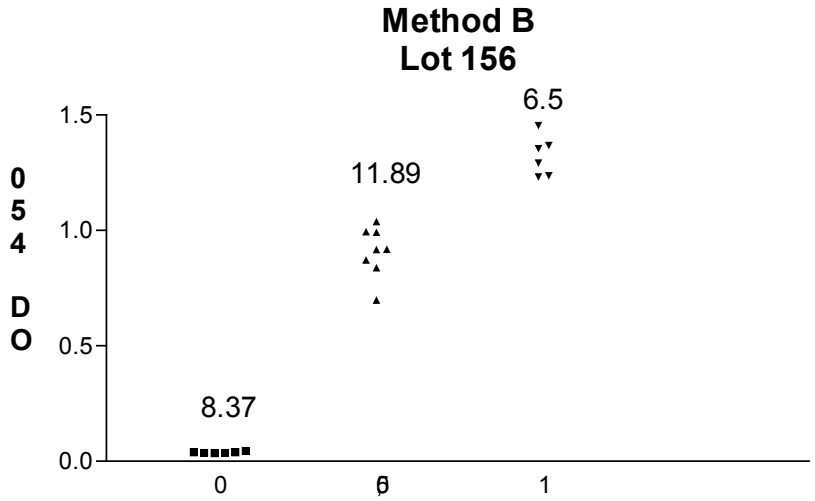


Fig 2b: Method B = KN CRYO WB/IL-1: blood frozen in the vapour phase of liquid nitrogen; Lot 156 8fold tested, 3 spikes

BRD: CRYO WB/IL-1 March, 2006

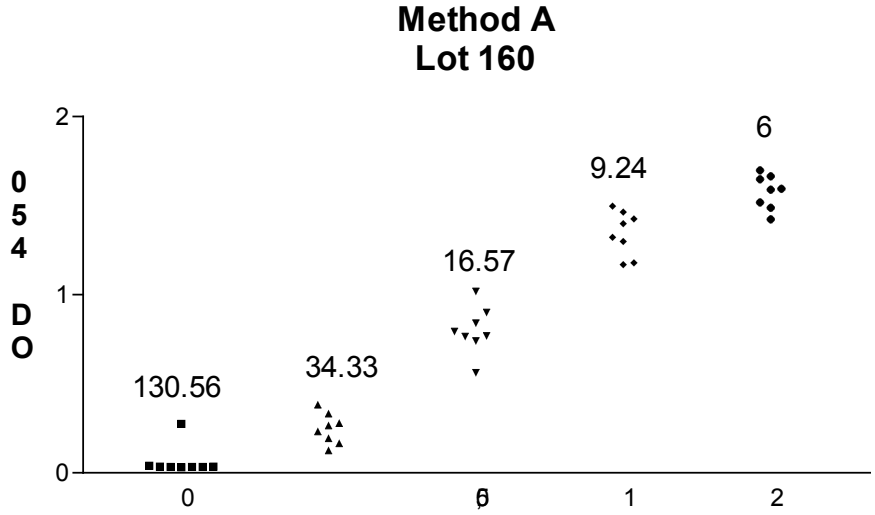


Fig 2c: Method A = CRYO WB/IL-1: blood frozen at -80°C;
Lot 160 8fold tested, 5 spikes

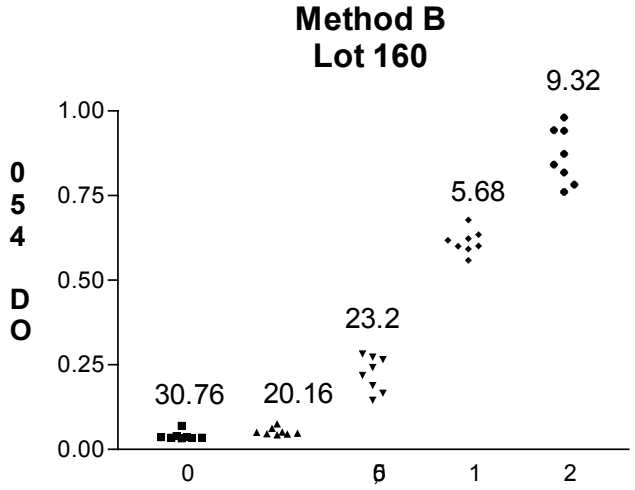


Fig 2d: Method B = KN CRYO WB/IL-1: blood frozen in the vapour phase of liquid nitrogen Lot 160 8fold tested, 5 spikes

BRD: CRYO WB/IL-1 March, 2006

3. Since KN CRYO/WB/IL-1 showed a higher variance and a lower limit of detection than the other two methods, a set of experiments was performed:

- reproducibility of saline and 0.5 EU/ml (corresponds to 1A in above Table 5.1.1)
- limit of detetction experiment (corresponds to 1B in above Table 5.1.1)

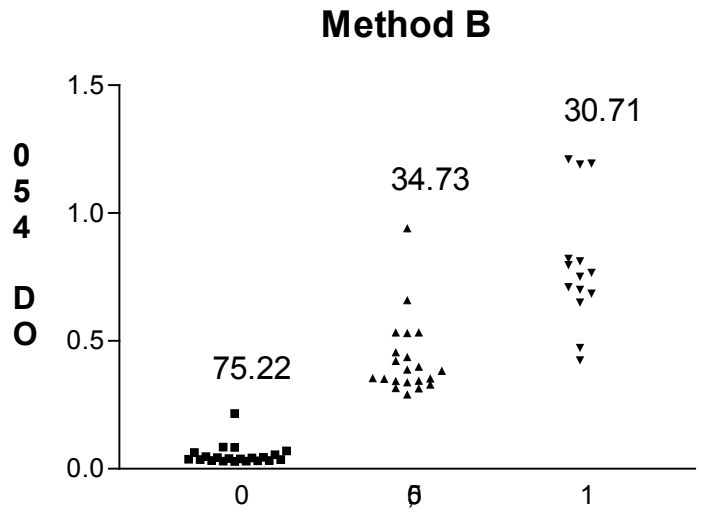
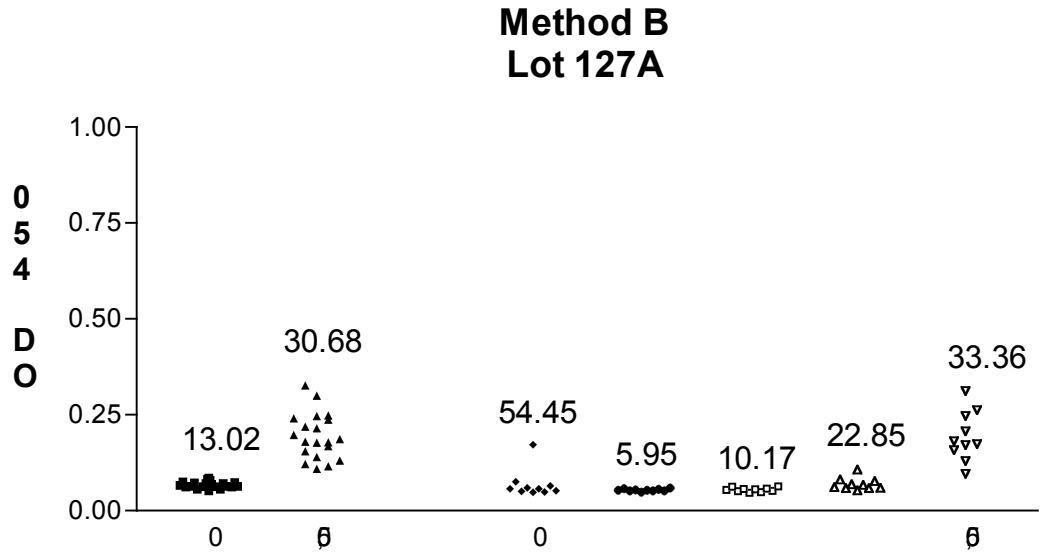


Fig 3: Method B = KN CRYO WB/IL-1: blood frozen in the vapour phase of liquid nitrogen; Variance of the method at spikes of 0 and 0.5 EU/ml: 18fold values; 1 EU/ml: 12fold values.

4. Due to the obvious lower sensitivity of KN CRYO WB/IL-1, an experiment was performed concerning the limit of detection according to 1B of the above table (Fig. 4a/b). This experiment was done twice with different pools of cryopreserved blood



BRD: CRYO WB/IL-1 March, 2006

Fig. 4a: Method B = KN CRYO WB/IL-1: Limit of detection of the frozen blood (nitrogen) Lot Number 127A. Stimulus: E. coli O113: H10:

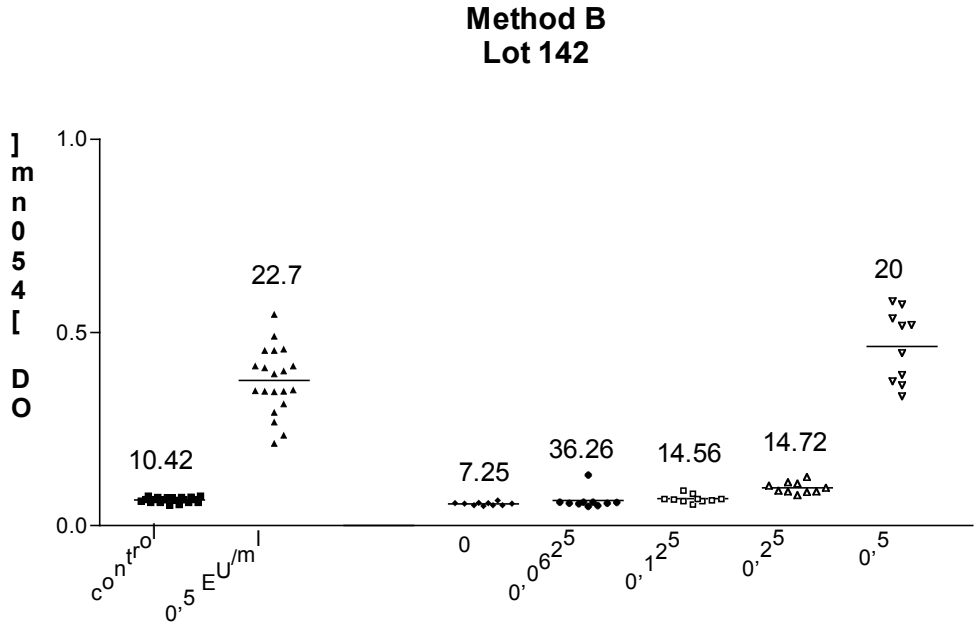


Fig. 4b: Method B = KN CRYO WB/IL-1: Limit of detection of the frozen blood (nitrogen), lot 142. Stimulus: E. coli O113: H10

All in all, these above experiments indicated no need for extensive studies concerning the intralaboratory performance and variability of the three methods. The only method that appeared problematic was KN CRYO WB/IL-1, whose characteristics were further explored with the above experiments. Nevertheless, KN CRYO WB/IL-1 proved to reliably retrieve the 0.5 EU/ml control as positive, despite a higher variability.

It was concluded, that the data from the different laboratories performing the methods in the prevalidation would suffice, taking the experiences of the former validation into account.

[This Page Intentionally Left Blank]

Validation of Biomedical Testing Methods

Human Whole Blood Pyrogen Test in 96-well Plates Using Fresh or Cryopreserved Blood

Standard Operating Procedure

Only the GLP/QA Unit is allowed to make copies of this document. Extra copies can be obtained on request from the GLP/QA Unit. Quality Documents are valid only if they are signed by the GLP/QA Unit and provided with a valid copy number.	**Copy number**

Date of Circulation: *July 02, 2004*
File Identifier: *SOP-WBT-KN2.v01*

JOINT RESEARCH CENTRE
EUROPEAN COMMISSION
Institute for Health and Consumer Protection
ECVAM

Standard Operating Procedure

Human Whole Blood Pyrogen Test in 96-well Plates Using Fresh or Cryopreserved Blood

Drafted by:	*Name*	*Stefanie Schindler*
	Date	*15/03/04*
	Signature	
Reviewed by:	*Name*	*Marlies Halder*
	Date	*29/06/04*
	Signature	
Approved by:	*Name*	
	Date	
	Signature	
Issued by:	*Name*	
	Date	
	Signature	

*Owner/Trainer: Signature: Date:

PAGE OF CHANGES

Date of change/ Date of draft:	Version-number:	Changed page(s):	Summary of the change(s):	Changed by/Sign.:

TABLE OF CONTENTS Page no.

1. INTRODUCTION 4
2. PURPOSE 5
3. SCOPE / LIMITATIONS 5
4. METHOD OUTLINE 5
5. DEFINITIONS / ABBREVIATIONS 6
6. MATERIALS 7
7. METHODS 9
8. HEALTH SAFETY AND ENVIRONMENT 14
9. ANNEX 15
10. REFERENCES 17

THIS SOP WAS AMENDED FOR THE CATCH-UP VALIDATION PHASE ONLY. IT DOES THEREFORE ONLY REPLACE THE PREVIOUS VERSION FOR THIS SERIES OF EXPERIMENTS.

1. INTRODUCTION

The whole blood pyrogen test (in vitro pyrogen test IPT) is a two-part assay for the detection of pyrogenic contamination. It involves incubation of the sample with human blood, followed by an enzyme immunoassay for the measurement of IL-1β.
A pyrogen is a substance that causes fever. Bacterial contaminations, which contain exogenous pyrogens, can be deadly. This problem is of great significance for drug safety.
Also, medical devices and biologically produced substances obtained from bacteria and other microorganisms may cause release of endogenous pyrogens (e.g., IL-1β).
Exogenous pyrogens include metabolic substances and cell-wall components of microorganisms. These substances are present during the "normal" course of an infectious disease. Infections by Gram-negative and Gram-positive bacteria are equal in frequency. Both of these bacterial types can activate the release of endogenous pyrogens, which cause fever through the thermoregulatory center in the brain. Although these reactions can occur during the "normal" course of an infectious disease, a deadly shock syndrome can occur in the worst case.
Due to these risks, product safety legislation demands rigorous quality checks for pyrogenic contamination of drugs and devices intended for parenteral use. For example, testing in rabbits for medical end products is required in Germany. Products in development and a few end products are allowed to be controlled by the Limulus assay. The first pyrogen assay, based on human whole blood stimulation by pyrogens, was developed by Hartung et al. (3,4).

2. PURPOSE

This assay simulates *in vitro* the normal human reaction to exogenous pyrogens. A few drops of human blood are mixed with the sample, and exogenous pyrogens in the sample are recognized by immunocompetent cells in the human blood.

These cells release IL-1β, which is measured by an integrated ELISA system.

3. SCOPE / LIMITATIONS

Limit of detection is ≤ 0.25 EEU/ml, not suitable for test samples interfering with blood cytokine release.

4. METHOD OUTLINE

The procedure has two parts:
1) Incubation of the sample with (diluted) human blood.
2) An enzyme immunoassay for the measurement of IL-1β.

Ad 1) Blood incubation
Diluted human whole blood is incubated for 10-24 hours together with saline or RPMI and the sample in a pyrogen-free microtiter plate and aliquots are taken for further examination.

Ad 2) Capture of Endogenous Pyrogens (ELISA procedure)
Samples (aliquots of whole blood stimulation) are distributed into the wells of a microplate which are coated with antibodies specific for IL-1β.

An enzyme-conjugated antibody against IL-1β is added. During a 90-minute incubation, a sandwich complex consisting of two antibodies and the IL-1β is formed. Unbound material is removed by a wash step.

A chromogenic substrate (3,3´,5,5´ -tetramethylbenzidine, TMB) reactive with the enzyme label is added. Color development is terminated by adding a stop solution after 30 minutes. The resulting color, read at 450 nm, is directly related to the IL-1β concentration. Bi-chromatic measurement with a 600-690 nm reference filter is recommended.

5. DEFINITIONS / ABBREVIATIONS

The following abbreviations are used in this work-book.

Ab	antibody
°C	degrees Celsius (Centigrade)
EC	endotoxin control
EEU	endotoxin equivalent unit
ELISA	Enzyme-Linked ImmunoSorbent Assay
EU	endotoxin unit of the international WHO standard
h	hour
HCl	hydrochloric acid
IL	interleukin
LPS	lipopolysaccharide (exogenous pyrogen from Gram-negative bacteria)
LTA	lipoteichoic acid (exogenous pyrogen from Gram-positive bacteria)
µl	microlitre
mg	milligram
min	minute
ml	millilitre
MTP	microtiter plate
MVD	maximum valid dilution
NaCl	sodium chloride, 0,9%
nm	nanometre
NPC	negative product control
PPC	positive product control
OD	optical density
rpm	rounds per minute
RT	room temperature
TMB	3,3´,5,5´-Tetramethylbenzidine

6. MATERIALS

6.1. Materials required and not provided

The components listed below are recommended, but equivalent devices may also be used: it is the users responsibility to validate the equivalence.
For all steps excluding the ELISA procedure sterile and pyrogen-free materials have to be used (e.g. tips, containers, solutions).

6.1.1 Materials for fresh blood incubation

Equipment
·Incubator (37°C + 5% CO_2)
·Multipette or adjustable 20 to 100 µl pipetters
·Multichannel pipettor, 8 or 12 channels
·Vortex mixer
·Laminar flow bench (recommended)

Consumables
·Heparinized tubes for blood sampling(Sarstedt S-MONOVETTE 7.5 ml, 15 IU/ml Li-Heparin)
·Sarstedt multifly needle set, pyrogen-free, for S-Monovette
·Non-pyrogenic 96-well polystyrene tissue culture microtiter plate, Falcon, Cat. No. 353072
·Sterile and pyrogen-free tips 20 µl and 100 µl
·Combitips for multipette, 1.0 ml and 0.5 ml
·Reservoir for saline
·Non-pyrogenic test tubes, preferably 12 or 15 ml centrifuge tubes from greiner bio-one or other qualified materials that can be used for preparing standards and diluting samples

6.1.2 Materials for incubation with cryopreserved blood

Equipment
·Incubator (37°C + 5% CO_2)
·Multipette or adjustable 20 to 100 µl pipetters
·Multichannel pipettor, 8 or 12 channels
·Vortex mixer
·Laminar flow bench (recommended)

Consumables
·Non-pyrogenic 96-well microtiter plate, Falcon, Cat No. 353072
·Sterile and pyrogen-free tips 20 µl and 100 µl
·Combitips for multipette, 1.0 ml and 0.5 ml
·Reservoirs for RPMI and saline
·Non-pyrogenic test tubes, preferably 12 or 15 ml centrifuge tubes from greiner bio-one or other qualified materials that can be used for preparing standards and diluting samples

6.1.3 Materials for ELISA procedure

Equipment
· Multichannel pipettor
· Microplate mixer
· Microplate washer (optional)
· Microplate reader capable of readings at 450 nm (optional reference filter in the range of 600-690 nm)
· A software package for facilitating data generation, analysis, reporting, and quality control

Consumables
· Graduated cylinder and plastic storage container for Buffered Wash Solution
· Tip-Tubs for reagent aspiration with Multichannel pipettor
· non-sterile pipette tips
· non-sterile deionized water

6.2. Materials Supplied in ELISA kit

Components supplied in that kit are *not* interchangeable with other lots of the same components.
IL-1β Ab-coated Microplate: One 96-well polystyrene microplate, packaged in a zip-lock foil bag, with desiccant. The plate consists of twelve strips mounted in a frame. Each strip includes eight anti-IL-1β Ab-coated wells. Additionally, individual strips can be separated from the frame to enable the repackaging and later use of all the wells of a kit. In this case, repackage the strips in the zip-lock foil with the desiccant, reseal the foil airtight and use the strips within 4 weeks. Well positions are indexed by a system of letters and numbers (A through H, 1 through 12) embossed on the left and top edges of the frame. Store refrigerated: stable at 2-8°C until the expiration date marked on the label.
Enzyme-Labeled Antibody : One amber vial containing 21 ml of liquid reagent, ready-to-use. The reagent contains horseradish peroxidase-labeled, affinity-purified, polyclonal anti-IL-1β antibodies, with preservative. Store refrigerated: stable at 2-8°C until the expiration date marked on the label. *Do not freeze.*
Saline: Three plastic vials, each containing pyrogen-free saline. This is intended for the dilution of fresh blood, samples and for reconstitution of the Endotoxin Control. Store refrigerated: stable at 2-8°C until the expiration date marked on the label. Use immediately after opening and discard unused volumes.
RPMI: One plastic vial, containing pyrogen-free RPMI. This is intended for dilution of cryopreserved blood. Store refrigerated: stable at 2-8°C until the expiration date marked on the label. Use immediately after opening and discard unused volumes.
TMB/Substrate Solution: Two amber vials, each containing 11 ml of a buffered reagent, ready-to-use. The reagent contains a hydrogen peroxide substrate and 3,3´,5,5´-tetramethylbenzidine (TMB). Store refrigerated and protected from light: stable at 2-8°C until the expiration date marked on the label. *Do not freeze*.
Buffered Wash Solution Concentrate: One vial containing 100 ml of a concentrated (10X) buffered saline solution, with surfactants and preservative. Using a transfer container, dilute the contents of the vial with **900 ml** distilled or deionized water for a total volume of **1000 ml**.
Store refrigerated: stable at 2-8°C until the expiration date marked on the label.

Stop Solution: One vial containing an acidic solution, for terminating the color reaction. The reagent is supplied ready-to-use. Handle with care, using safety gloves and eye protection. Store refrigerated: stable at 2-8°C until the expiration date marked on the label.

Additionally supplied materials

Endotoxin Control: One vial of an endotoxin control. The control is supplied lyophilized. Before use, reconstitute control vial with pyrogen-free distilled water. Prepare serial dilutions in saline (see 7. Methods). Mix by vortexing. After preparation, the stock solution can be stored (see 7. Methods).
PPC (Positive Product Control): one glass vial containing 1.05 ng/ml Endotoxin. Store at 2-8°C and use according to 7. Methods.

7. METHODS

7.1. Fresh Blood Incubation (Method 7 A)

Blood Collection
Collect blood by venipuncture into heparinized tubes. The blood collection system must be pyrogen-free. The procedure calls for 20 µl of heparinized whole blood per well. The blood can be stored in the collection tube at room temperature (15-28°C) for 4 hours. Incubation of the sample should be started within this time. Prior to use, gently invert the collection tube once or twice. **Do not vortex.**

Note:

1 Blood donors are to describe themselves as in good health and not in need of medication for the last two weeks.
2 Each assay should include the Endotoxin Controls and the saline control in quadruplicate.
3 Use disposable tip pipets to avoid contamination of reagents and samples.
4 During ELISA procedure, the wells should be washed carefully.
5 The test samples should be done in quadruplicate.
6 The contents of the wells must be decanted or aspirated completely before pipetting wash solution.
7 Deviations from the procedure (incubation time/temperature) may cause erroneous results. The ELISA procedure should be run without interruption. Diluted samples should be tested within an hour.

7.2. Blood incubation with cryopreserved blood (Method 7 B, Method 7 C)

Blood frozen according to the Konstanz method has to be stored in the vapour phase of liquid nitrogen
Blood frozen according to the PEI method can be kept at –80°C or in the vapour phase of liquid nitrogen; for longer storage, please transfer the vials into the vapour phase of liquid nitrogen.

Thawing procedure
Take the required number of aliquots out of the vapor phase of liquid nitrogen/the freezer and leave the blood to thaw in the incubator at 37°C for 15 minutes. After this time, dry the condensed water off the vials using a paper cloth. Preferably under a laminar-flow bench, unscrew the vials and pool the blood in a polypropylene centrifuge tube. Gently invert the tube once or twice to achieve complete mixing. **Do not vortex.**

Storage of the substances

- please keep all substances and spikes at 4°C

Spiking of the substances

Part 1)
5 blinded spikes have been sent out by Konstanz
They are bearing a code for

a) the respective drug
b) a random blinding number

- please pipet 500 µl of the respective substance into a test tube
- vortex the respective vial with the blinded spike for about 5 seconds
- add 25 µl of the spike to the substance and vortex for another 5 seconds
- perform the dilutions according to the instructions below

In case of little substance, the amounts may be reduced to 250 µl of substance + 12.5 µl of spike.

Dilution of the substances

- for dilution, please use either 12 ml or 15 ml tubes from greiner bio-one
- each substance has to be vortexed for about 5 seconds immediately before performing Step 3 of the Whole Blood Stimulation.

Substance 1: Glucose 5%
Maximum valid dilution: 1:70; add 50 µl of substance to 3450 µl of saline

Substance 2: EtOH 13%
Maximum valid dilution: 1:35 ; add 100 µl of substance to 3400 µl of saline

Substance 3: MCP
Maximum valid dilution: 1:350; add 10 µl of substance to 3490 µl of saline

Substance 4: Syntocin
Maximum valid dilution: 1:700: add 5 µl of substance to 3495 µl of saline

Substance 5: Binotal
Maximum valid dilution: 1:140; add 25 µl of substance to 3475 µl of saline

Substance 6: Fenistil
Maximum valid dilution: 1:175; add 20 µl of substance to 3480 µl of saline

Substance 7: Sostril
Maximum valid dilution: 1:140; add 25 µl of substance to 3475 µl of saline

Substance 8: Beloc
Maximum valid dilution: 1:140; add 25 µl of substance to 3475 µl of saline

Substance 9: Drug A
Maximum valid dilution: 1:35; add 100 µl of substance to 3400 µl of saline

Substance 10: Drug B
Maximum valid dilution: 1:70; add 50 µl of subsubstance to 3450 µl of saline

Part 2)
(unblinded)

- Positive Product Control (PPC)
dilute the respective substance according to the instructions above
vortex for about 5 seconds
pipet 500 µl of the diluted substance into a pyrogen-free tube
add 25 µl of the unblinded PPC-LPS spike

- Negative Product Control (NPC)
dilute the respective substance according to the instructions above
vortex for about 5 seconds
pipet 500 µl of the diluted substance into a pyrogen-free tube
add 25 µl of saline

Endotoxin dilution for the Dose-Response Curve

IPT assays must include the 0.5 EU/ml + saline control in quadruplicate.

Dissolve the contents of the vial containing O113 provided by NIBSC with 5 ml of pyrogen-free distilled water yielding a stock solution of 2000 EU/ml. After reconstitution of the lyophilisate, vortex the stock solution according to the Certificate of Analysis. Vortex all dilutions prior to use for 5 seconds.

EC = Endotoxin Control, for use in the assay.

Solution	amount added to saline	Volume of saline	Resulting solution
Stock (2000 EU/ml)	100 µl	900 µl	200 EU/ml
200 EU/ml	100 µl	900µl	20 EU/ml
20 EU/ml	100 µl	900 µl	2 EU/ml
2 EU/ml	500 µl	500 µl	1 EU/ml (EC)
1 EU/ml	500 µl	500 µl	0,5 EU/ml (EC)

The stock solution of the Endotoxin Standard may be aliquoted (e.g. 100 µl aliquots) and kept at –20 °C for up to 6 months. Do not store the O113 at –80°C.

Whole Blood Stimulation, fresh blood (Method 7 A)

Perform incubation of blood samples in a microtiter plate.Preferably, use a laminar-flow bench. All consumables and solutions have to be sterile and pyrogen-free.

Step 1: Draw up an incubation plan according to the template below

Step 2: Add **200 µl** saline into each well.

Step 3: Add **20 µl** of Endotoxin Controls and negative saline control or samples in quadruplicate into the respective wells according to the prepared incubation plan.

Step 4: Add **20 µl** of donor blood, mixed by gentle inversion, into each well.

Step 5: Mix the contents of the wells thoroughly by gently aspiring and dispensing them 5 times, using a multichannel pipettor, changing the tips between each row in order to avoid cross-contamination.

Step 6: Put the lid on the plate and place the plate in an incubator at 37°C + 5%CO_2 for 10-24 hours.

Step 7: When transferring the stimulation aliquots onto the ELISA plate, mix the contents of the wells thoroughly by gently aspiring and dispensing them 3 times, using a multichannel pipettor, changing the tips between each row in order to avoid cross-contamination.

The aliquots can be tested immediately by the ELISA System or may be stored at -20°C or –80°C for testing at a later time. After transfer onto the ELISA plate, keep the remaining stimulation aliquots in the incubation plate at 20/-80°C for eventual repetition of the ELISA procedure (see Minimum assay suitability requirements).

Whole Blood Stimulation, cryopreserved blood (Method PEI = Method 7 B)

Perform incubation of blood samples in a microtiter plate. Preferably, use a laminar-flow bench**.** All consumables and solutions have to be sterile and pyrogen-free.

Step 1: Draw up an incubation plan according to the template below.

Step 2: Add **180 µl** RPMI into each well.

Step 3: Add **20 µl** of Endotoxin Controls and negative saline control or samples in quadruplicate into the respective wells according to the prepared incubation plan.

Step 4: Add **40 µl** of donor blood, mixed by gentle inversion, into each well.

Step 5: Mix the contents of the wells thoroughly by gently aspiring and dispensing them 5 times, using a multichannel pipettor, changing the tips between each row in order to avoid cross-contamination.

Step 6: Put the lid on the plate and place the plate in an incubator at 37°C + 5%CO_2 for 10-24 hours.

Step 7: When transferring the stimulation aliquots onto the ELISA plate, mix the contents of the wells thoroughly by gently aspiring and dispensing them 3 times, using a multichannel pipettor, changing the tips between each row in order to avoid cross-contamination.

The aliquots can be tested immediately by the ELISA System or may be stored at -20°C or –80°C for testing at a later time. After transfer onto the ELISA plate, keep the remaining stimulation aliquots in the incubation plate at 20/-80°C for eventual repetition of the ELISA procedure (see Minimum assay suitability requirements).

Whole Blood Stimulation, cryopreserved blood (Method Konstanz= Method 7 C)

Perform incubation of blood samples in a microtiter plate. Preferably, use a laminar-flow bench**.** All consumables and solutions have to be sterile and pyrogen-free.

Step 1: Draw up an incubation plan according to the template below

Step 2: Add **200 µl** RPMI into each well.

Step 3: Add **20 µl** of Endotoxin Controls and negative saline control or samples in quadruplicate into the respective wells according to the prepared incubation plan.

Step 4: Add **20 µl** of donor blood, mixed by gentle inversion, into each well.

Step 5: Mix the contents of the wells thoroughly by gently aspiring and dispensing them 5 times, using a multichannel pipettor, changing the tips between each row in order to avoid cross-contamination.

Step 6: Put the lid on the plate and place the plate in an incubator at 37°C + 5% CO_2 for 10-24 hours.

Step 7: Take the plate out of the incubator and freeze it at –20 or –80°C until the contents of the wells are completely frozen. After this, thaw the plate at room temperature or in a water bath at no more than 37°C.

Step 8: When transferring the stimulation aliquots onto the ELISA plate, mix the contents of the wells thoroughly by gently aspiring and dispensing them 3 times, using a multichannel pipettor, changing the tips between each row in order to avoid cross-contamination.

The aliquots can be tested immediately by the ELISA System or may be stored at -20°C or –80°C for testing at a later time. After transfer onto the ELISA plate, keep the remaining stimulation aliquots in the incubation plate at 20/-80°C for eventual repetition of the ELISA procedure (see Minimum assay suitability requirements).

7.2. ELISA Procedure

Remove the ELISA kit from the refrigerator at least 30 minutes before use. All components must be at room temperature (15-28°C). The ELISA is carried out at room temperature.

1 Sample distribution: see Microplate Template below.

A	NPC (A)	NPC (A)	PPC (A)	PPC (A)	PPC (A)	PPC (A)	1 (B)	1 (B)	1 (B)	1 (B)	2 (C)	2 (C)
B	NPC (A)	NPC (A)	1 (A)	1 (A)	1 (A)	1 (A)	2 (B)	2 (B)	2 (B)	2 (B)	2 (C)	2 (C)
C	EC 1,0	EC 1,0	2 (A)	2 (A)	2 (A)	2 (A)	3 (B)	3 (B)	3 (B)	3 (B)	3 (C)	3 (C)
D	EC 1,0	EC 1,0	3 (A)	3 (A)	3 (A)	3 (A)	4 (B)	4 (B)	4 (B)	4 (B)	3 (C)	3 (C)
E	EC 0,5	EC 0,5	4 (A)	4 (A)	4 (A)	4 (A)	5 (B)	5 (B)	5 (B)	5 (B)	4 (C)	4 (C)
F	EC 0,5	EC 0,5	5 (A)	5 (A)	5 (A)	5 (A)	NPC (C)	NPC (C)	NPC (C)	NPC (C)	4 (C)	4 (C)
G	saline	saline	NPC (B)	NPC (B)	NPC (B)	NPC (B)	PPC (C)	PPC (C)	PPC (C)	PPC (C)	5 (C)	5 (C)
H	saline	saline	PPC (B)	PPC (B)	PPC (B)	PPC (B)	1 (C)	1 (C)	1 (C)	1 (C)	5 (C)	5 (C)

A, B, C : e.g. Substances 1, 2, 3
1-5 : blinded spikes 1-5
EC : Endotoxin Control

NPC: negative product control
PPC: positive product control

2 Add **100 µl** Enzyme-Labeled Antibody to every well.

3 Within 10 minutes, pipet **100 µl** of whole blood stimulations of Endotoxin Controls, those of the negative saline control and of the samples into the wells prepared. During transfer, resuspend the contents of the wells of the incubation plate by aspiring and dispensing them 3 times.

Use a disposable-tip micropipet for the samples, changing the tip between each sample and control, to avoid contaminations.

4 Seal the plate with the adhesive foil provided in the kit.

5 Mix for **90 minutes** on a microplate mixer at 350-400 rpm.

6 Decant, then wash. Wash each well 4 times with 250-300 µl Buffered Wash Solution.

If this step is performed manually, remove as much moisture as possible during the decanting by inverting the washed microplate and tapping out the residual washing buffer on blotting paper or a paper towel, being careful not to dislodge the strips from the frame. Perform this step before adding the TMB.

7 Add **200 µl** of TMB/Substrate Solution to every well.

8 Incubate without shaking for **15 minutes** in the dark. Reduce incubation time if necessary (see Minimum assay suitability requirements).

9 Add **50 µl** of Stop Solution to every well.

*Tapping the plate **gently** after the addition of Stop Solution will aid mixing and improve precision. The Stop Solution is acidic.*
Handle carefully, and use safety gloves and eye protection.
10 Read at 450 nm, **within 15 minutes** of adding Stop Solution. Bi-chromatic measurement with a reference wavelength of 600-690 nm is recommended.

MINIMUM ASSAY SUITABILITY REQUIREMENTS

The assay should be considered acceptable only if the following minimum criteria are met:

The mean OD of the 0.5 EU/ml endotoxin control is at 1.6 times the mean OD of the negative saline control or greater.

The mean OD of the PPC is at 1.6 times the mean OD of the NPC or greater.

The mean OD of the PPC has to be in the 50-200 % range of the mean OD of the 0.5 EU/ml endotoxin control.

The mean OD of the negative saline control is at 100 mOD or lower.

If one OD value of the of the 1.0 EU/ml Endotoxin Control is > Max, the ELISA procedure may be repeated, reducing the incubation time (Step 8 of 7.2. ELISA Procedure).

8. HEALTH SAFETY AND ENVIRONMENT

· For *in vitro* use only.
· Do not use reagents beyond their expiration dates.

Bio-Safety
Human blood has to be considered infectious and handled accordingly. When handling nitrogen and the unopened vials of cryopreserved blood, wear protective eyewear. Wear gloves when performing incubations.

Stop Solution and TMB/Substrate Solution
Avoid contact with the Stop Solution, which is acidic. Wear gloves and eye protection. If this reagent comes into contact with skin, wash thoroughly with water and seek medical attention, if necessary. The reagent is corrosive; therefore, the instrument employed to dispense it should be thoroughly cleaned after use. The TMB/Substrate Solution contains peroxide. Since peroxides are strong oxidizing agents, avoid all bodily contact with the TMB/Substrate Solution

9. ANNEX (Pipetting scheme for the whole blood assay)

Part 1a: Whole blood stimulation, fresh blood (all values in µl) (Test 7.1)

well account	Stimulation sample	saline	Endotoxin Control (0.5 – 1.0 EU/ml)	Test sample	Donor blood		
4	Endotoxin Control (0.5 – 1.0 EU/ml)	200	20	-	20	Mix the samples. Incubate overnight at 37°C + 5% CO_2	Mix the samples. Test immediately with the ELISA system or store at -20/-80 °C.
4	Blank (0)	220	-	-	20		
4	Test samples (1-10)	200	-	20	20		

Part 1b: Whole blood stimulation, cryopreserved blood, PEI method (all values in µl) (Test 7.2)

well account	Stimulation sample	RPMI	saline	Endotoxin Control (0.5 – 1.0 EU/ml)	Test sample	Donor blood		
4	Endotoxin Control (0.5 – 1.0 EU/ml)	180	-	20	-	40	Mix the samples. Incubate overnight at 37°C + 5% CO_2	Mix the samples. Test immediately with the ELISA system or store at -20/-80 °C.
4	Blank (0)	180	20	-	-	40		
4	Test samples (1-10)	180	-	-	20	40		

Part 1c: Whole blood stimulation, cryopreserved blood, Konstanz method (all values in µl) (Test 7.3)

well account	Stimulation sample	RPMI	saline	Endotoxin Control (0.5 – 1.0 EU/ml)	Test sample	Donor blood		
4	Endotoxin Control (0.5 – 1.0 EU/ml)	200	-	20	-	20	Mix the samples. Incubate overnight at 37°C + 5% CO_2	Mix the samples. Test immediately with the ELISA system or store at -20/-80 °C.
4	Blank (0)	200	20	-	-	20		
4	Test samples (1-10)	200	-	-	20	20		

Part 2: ELISA procedure (all values in µl)

Supernatants from Stimulation	Enzyme-labeled Antibody		Substrate		Stop solution	
100	100	Incubate 90 min at RT on a plate mixer at 350-400 rpm. Decant.Wash 4 times with 300 µl Buffered Wash Solution	200	Incubate 15 min at RT in a dark place	50	Read at 450 nm (600-690 nm reference wave-length recom-mended)

10. REFERENCES

1 Hartung T., Wendel A. : Detection of pyrogens using human whole blood. In Vitro Toxicology; 9(4): 353-59.

2 Fennrich S., Fischer M., Hartung T., Lexa P., Montag-Lessing T., Sonntag H.-G., Weigand M. und Wendel A. : Detection of endotoxins and other pyrogens using human whole blood . Dev Biol Stand.Basel , 1999, 101:131-39.

3 Hartung T., Aaberge I., Berthold S., Carlin G., Charton E., Coecke S., Fennrich S., Fischer M., Gommer M., Halder M., Haslov K., Montag-Lessing T., Poole S., Schechtman L., Wendel A. und Werner-Felmayer, G. : ECVAM workshop on novel pyrogen tests based on the human fever reaction. ATLA 2001, 29:99-123.

4 Morath S., Geyer A., Hartung T. : Structure-function relationship of cytokine induction by lipoteichoic acid from Staphylococcus aureus. J. Exp. Med., 2001, 193:393-397.

5 Bonenberger J., Diekmann W., Fennrich S., Fischer M., Friedrich A., Hansper M., Hartung T., Jahnke M., Löwer J., Montag-Lessing T., Petri E., Sonntag H.-G., Weigand M., Wendel A. und Zucker B. : Pyrogentestung mit Vollblut-Zusammenfassung eines Status-Workshops am Paul-Ehrlich-Institut, Langen, am 22.11.99. Bundesgesundheitsbl-Gesundheitsforsch-Gesundheitsschutz 2000, 43:525-533.

6 Petri E., van de Ploeg A., Habermaier B. and Fennrich S.: Improved detection of pyrogenic substances on polymer surfaces with an ex vivo human whole-blood assay in comparison to the Limulus amoebocyte lysate test. In: Balls M., van Zeller A.-M. and Halder M.: Progress in the reduction, refinement and replacement of animal experimentation. Elsevier 2000, 339-345.

7 Fennrich S., Wendel A. and Hartung T.: New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 1999, 16:146-149.

Appendix A3

The Human WB/IL-6 *In Vitro* Pyrogen Test

ECVAM Background Review Document (March 2006) A-239
ECVAM Standard Operating Procedure (July 2002) A-307

[This Page Intentionally Left Blank]

BRD: WB/IL-6 March, 2006

THE HUMAN WHOLE BLOOD/IL-6 IN VITRO PYROGEN TEST (WB/IL-6)

BRD: WB/IL-6 March, 2006

Contents

1 RATIONALE FOR THE PROPOSED TEST METHOD 4

1.1 INTRODUCTION 4
1.2 REGULATORY RATIONALE AND APPLICABILITY 5
1.3 SCIENTIFIC BASIS FOR THE PROPOSED TEST METHOD 6

2 TEST METHOD PROTOCOL COMPONENTS 8

2.1 OVERVIEW OF TEST METHOD 8
2.2 RATIONAL FOR SELECTED TEST COMPONENTS 9
2.3 BASIS FOR SELECTION OF THIS TEST METHOD 13
2.4 PROPRIETARY COMPONENTS 13
2.5 REPLICATES 13
2.6 MODIFICATIONS APPLIED AFTER VALIDATION 14
2.7 DIFFERENCES WITH SIMILAR TEST METHODS 14

3 SUBSTANCES USED FOR VALIDATION 15

3.1 SELECTION OF SUBSTANCES USED 15
3.2 NUMBER OF SUBSTANCES 15
3.3 DESCRIPTION OF SUBSTANCES USED 16
3.4 SAMPLE CODING PROCEDURE 16
3.5 RECOMMENDED REFERENCE CHEMICALS 17

4 *IN VIVO* REFERENCE DATA ON ACCURACY 18

4.1 TEST PROTOCOL *IN VIVO* REFERENCE TEST METHOD 18
4.2 ACCURACY 18
4.3 ORIGINAL RECORDS 19
4.4 QUALITY OF DATA 19
4.5 TOXICOLOGY 20
4.6 BACKGROUND ON ASSAY PERFORMANCE 20

5 TEST METHOD DATA AND RESULTS 21

5.1 TEST METHOD PROTOCOL 21
5.2 ACCURACY AND RELIABILITY 23
5.3 STATISTICS 28
5.4 TABULATED RESULTS 32
5.5 CODING OF DATA 33
5.6 CIRCUMSTANCES 34
5.7 OTHER DATA AVAILABLE 34

6 TEST METHOD ACCURACY 35

6.1 ACCURACY 35
6.2 CONCORDANCY TO *IN VIVO* REFERENCE METHOD 37
6.3 COMPARISON WITH REFERENCE METHODS 37
6.4 STRENGTH AND LIMITATIONS 37
6.5 DATA INTERPRETATION 38
6.6 COMPARISON TO OTHER METHODS 38

7 TEST METHOD RELIABILITY (REPEATABILITY/REPRODUCIBILITY) 39

7.1 SELECTION OF SUBSTANCES 39
7.2 RESULTS 39
7.3 HISTORICAL DATA 44
7.4 COMPARISON TO OTHER METHODS 44

8 TEST METHOD DATA QUALITY 45

8.1 CONFORMITY 45
8.2 AUDITS 45
8.3 DEVIATIONS 45
8.4 RAW DATA 45

9 OTHER SCIENTIFIC REPORTS AND REVIEWS 46

9.1 SUMMARY 46
9.2 DISCUSSION 48
9.3 RESULTS OF SIMILAR VALIDATED METHOD 48

10 ANIMAL WELFARE CONSIDERATIONS (REFINEMENT, REDUCTION, AND REPLACEMENT) 51

10.1 DIMINISH ANIMAL USE 51
10.2 CONTINUATION OF ANIMAL USE 51

11 PRACTICAL CONSIDERATIONS 53

11.1 TRANSFERABILITY 53
11.2 TRAINING 53
11.3 COST CONSIDERATIONS 54
11.4 TIME CONSIDERATIONS 54

12 REFERENCES 55

13 SUPPORTING MATERIALS (APPENDICES) 58

13.1 STANDARD OPERATING PROCEDURE (SOP) OF THE PROPOSED METHOD 58
13.2 STANDARD OPERATING PROCEDURE (SOP) OF THE REFERENCE METHOD 58
13.3 PUBLICATIONS 58
13.4 ORIGINAL DATA 61
13.5 PERFORMANCE STANDARDS 61

Appendix A – Method protocol(s) and trial plan(s)

Appendix B – Hardcopies of relevant publications

Appendix C – List of abbreviations and definitions

BRD: WB/IL-6 March, 2006

1 Rationale for the Proposed Test Method

1.1 Introduction

1.1.1. *Describe the historical background for the proposed test method, from original concept to present. This should include the rationale for its development, an overview of prior development and validation activities, and, if applicable, the extent to which the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards.*

Pyrogens, a chemically heterogeneous group of hyperthermia- or fever-inducing compounds, derive from bacteria, viruses, fungi or from the host himself reacting to microbial products during an immune response by producing endogenous pyrogens such as prostaglandins and the pro-inflammatory cytokines interleukin-1 (IL-1), interleukin-6 (IL-6) and tumour necrosis factor-α (TNF-α) (Dinarello, 1999). Depending on the type and amount of pyrogen challenge and the sensitivity of an individual, even life-threatening shock-like conditions can be provoked. To assure quality and safety of any pharmaceutical product for parenteral application in humans, pyrogen testing is therefore imperative.

Depending on the medicinal product, one of two animal-based pyrogen tests is currently prescribed by the respective Pharmacopoeias, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET). For the rabbit pyrogen test, sterile test substances are injected intravenously to rabbits and any rise in body temperature is assessed. This *in vivo* test detects various pyrogens but not alone the fact that large numbers of animals are required to identify a few batches of pyrogen-containing samples argues against its use. In the past two decades, the declared intention to refine, reduce and replace animal testing, has lowered rabbit pyrogen testing by 80% by allowing to use the BET as an *in vitro* alternative pyrogen test for certain medicinal products (Cooper et al, 1971).

Bacterial endotoxin, comprising largely lipopolysaccharide (LPS) from the cell wall of Gram-negative bacteria that stimulates monocytes/macrophages via interaction with CD14 and toll-like receptor 4 (TLR4) (Beutler and Rietschel, 2003), is the pyrogen of major concern to the pharmaceutical industry due to its ubiquitous sources, its stability and its high pyrogenicity. With the BET, endotoxin is detected by its capacity to coagulate the amoebocyte lysate from the haemolymph of the American horseshoe crab, *Limulus polyphemus*, a principle recognised some 40 years ago (Levin and Bang, 1964). In the US, *Limulus* crabs are generally released into nature after drawing about 20% of their blood and therefore most of these animals survive. However, the procedure still causes mortality of about 30,000 horseshoe crabs per year, which adds to even more efficient threats of the horseshoe crab population like its use as bait for fisheries, habitat loss and pollution (http://www.horseshoecrab.org/).

As with the rabbit test the general problem of translation of the test results to the human fever reaction persists. Moreover, although being highly sensitive, the failure of the BET to detect non-endotoxin pyrogens as well as its susceptibility to interference by e.g. high protein or lipid levels of test substances or by glucans impedes full replacement of the rabbit pyrogen test. Hence, hundreds-of-thousands rabbits per year are still used for pyrogen testing.

A test system that combines the high sensitivity and *in vitro* performance of the BET test with the wide range of pyrogens detectable by the rabbit pyrogen test is therefore required in order to close the current testing gap for pyrogens and to avoid animal-based tests. With this intention and due to improved understanding of the human fever reaction (Dinarello, 1999), test systems based on *in vitro* activation of human monocytes were developed. First efforts date back about 20 years, when peripheral blood mononuclear cells (PBMC) were used to detect endotoxin by monitoring the release of pyrogenic cytokines (Duff and Atkins, 1982; Dinarello et al 1984). Meanwhile, a number of different test systems, using either whole blood, peripheral blood mononuclear cells (PBMCs) or the monocytoid cell lines MONO MAC 6 (MM6) or THP-1 as a source for human monocytes and various read-outs were established (Poole et al., 1988; Ziegler et al, 1988; Tsuchiya et al, 1980; Hartung and Wendel, 1996; Hartung et al, 2001; Poole et al, 2003). These test systems were validated with the aim of developing a tool for formal inclusion into Pharmacopoeias, an important basis for implementing novel alternative pyrogen tests for product-specific validation.

1.1.2 *Summarize and provide the results of any peer review conducted to date and summarize any ongoing or planned reviews.*
All of the five methods are currently under peer review of the ECVAM Scientific Advisory Committee.

1.1.3 *Clearly indicate any confidential information associated with the test method; however, the inclusion of confidential information is discouraged.*
This document does not contain any confidential information.

1.2 Regulatory rationale and applicability

1.2.1 *Describe the current regulatory testing requirement(s) for which the proposed test method is applicable.*
To assure quality and safety of pharmaceutical products for parenteral application in humans, pyrogen testing is imperative. Depending on the drug, one of two pyrogen tests is currently prescribed by the European Pharmacopoeia, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET), and other national and international guidelines.

1.2.2 *Describe the intended regulatory use(s) (e.g., screen, substitute, replacement, or adjunct) of the proposed test method and how it will be used to substitute, replace, or complement any existing regulatory testing requirement(s).*
Dependent on the product and the presence of relevant clinical data on unexpected pyrogenicity of clinical lots, the proposed test method may be an alternative method for pyrogen testing, thus substituting the rabbit pyrogen test or the BET. In certain cases, the proposed test method may function as a supplementary test method to assess compliance to the licensing BRD.
In case the proposed test method is an alternative for pyrogenicity testing, a thorough cross-validation between the proposed test method and the original method for the specific medicinal product is warranted. In case the proposed test method is an adjunctive test to screen for (unexpected) pyrogenic lots, alert and alarm limits may be established based on consistency of production lots or (preferably) based on actual clinical data.

1.2.3 *Where applicable, discuss the similarities and differences in the endpoint measured in the proposed test method and the currently used in vivo reference test method and, if appropriate, between the proposed test method and a comparable validated test method with established performance standards.*
The current *in vivo* method (rabbit test), as described in the pharmacopoeia, and the proposed *in vitro* test method each determine very different end-points, though the biochemical origins of the response are similar.
The *in vivo* method more resembles a black box, and determines the total rise in body temperature (fever induction) of the animals subjected to the medicinal product, as a result of pyrogens (if any) present in the product.
The proposed test method WB/IL-6 is an *in vitro* model for the fever response mechanism. It determines the release of cytokines by monocytoid cells into the culture medium upon the interaction of pyrogens and specific receptors on the monocytoid cells. It is these cytokines that trigger the fever response *in vivo*.
Main differences between the *in vivo* and *in vitro* methods are that the latter is quantitative and uses cells of human origin, thus better reflecting the physiological situation.

1.2.4 *Describe how the proposed test method fits into the overall strategy of hazard or safety assessment. If a component of a tiered assessment process, indicate the weight that should be applied relative to other measures.*
The proposed test method WB/IL-6 may be applied for those medicinal products for which the rabbit test is the only or most reliable method for pyrogenicity testing, since a) the medicinal product is not compatible with the BET or b) the medicinal product contains pyrogens other than Gram-negative endotoxin.
Limit concentrations for pyrogens are established based on consistency lots or actual clinical data or, in the case of endotoxin the ELC as defined for many medicinal products.

1.3 Scientific basis for the proposed test method

1.3.1 *Describe the purpose and mechanistic basis of the proposed test method.*
The proposed *in vitro* method is intended to determine the presence of pyrogens in medicinal products for parenteral use. The proposed test method is an *in vitro* model of the human fever response. It determines the release of cytokines upon the interaction of pyrogens and specific Toll-like receptors on the monocytoid cells (Beutler and Rietschel, 2003). It is these cytokines that trigger the fever response *in vivo*.

1.3.2 *Describe what is known and not known about the similarities and differences of modes and mechanisms of action in the proposed test method as compared to the species of interest (e.g., humans for human health-related toxicity testing).*
An important feature of the proposed test method is that it is based upon the use of monocytoid cells of human origin. It therefore by definition resembles more closely the actual response of humans. The two other test methods make use of either crustaceans (BET) or rabbits, both species more or less distinct from the human species. The response of humans, horseshoe crabs and rabbits toward Gram-negative endotoxin has been studied extensively and the methods appear equivalent for this particular pyrogen

(Cooper et al 1971; Greisman and Hornick, 1969). However, there are documented cases of medicinal products and specified pyrogenic substances that yield false-positive or false-negative results in either test method. Since the proposed test method is based on human cells, it may therefore predict more accurately the pyrogenicity of such substances in humans.

1.3.3 *Describe the intended range of substances amenable to the proposed test method and/or the limits of the proposed test method according to chemical class or physicochemical factors.*

The proposed test method is intended for the assessment of pyrogens in all parenteral medicinal products for human use, chemical or biological and including raw materials, bulk ingredients and excipients. Use of the proposed test method in testing environmental samples or medicinal products is suggested and may be feasible, but substantiating data are as yet limited or absent.

2 Test Method Protocol Components

2.1 Overview of test method.

Provide an overview of how the proposed test method is conducted. If appropriate, this would include the extent to which the protocol for the proposed test method adheres to established performance standards.

A highly detailed method protocol describing the proposed test method WB/IL-6 (*Detailed protocol WB/Il-6 In vitro test for pyrogen/endotoxin using human whole blood 22 07 02*) is included in Appendix A of this background review document (BRD).

The WB/IL-6 test method is a two-part assay for the detection of pyrogenic contamination. The test protocol itself can be divided into the following two parts:

1. Incubation of the sample with (diluted) human blood.
2. An enzyme linked immunoassay (ELISA) for the measurement of IL-6.

Ad 1.
Human whole blood is collected by venipuncture into tubes for blood sampling and heparinized. Freshly collected (< 4 hours) heparinized human whole blood is incubated overnight (16-24 hours) together with saline and the sample of interest in sterile and pyrogen-free reaction tube. The supernatant is subsequently collected for further examination.

Ad 2.
Samples (supernatants of blood stimulation) are distributed into the wells of a microtiterplate which are coated with monoclonal antibodies specific for IL-6.
An enzyme-conjugated polyclonal antibody against IL-6 is added. During a subsequent incubation, a sandwich complex consisting of two antibodies and the IL-6 is formed. Unbound material is removed by a wash step.
A chromogenic substrate reactive with the enzyme label is added. Color development is terminated by adding a stop solution. The resulting color, read at the appropriate wavelength (substrate-dependent), is directly related to the IL-6 concentration.
The IL-6 ELISA used throughout this study is an in-house ELISA, developed by Novartis, in which the IL-6 calibrant is calibrated against the IS for IL-6 (89/548).

The WHO-LPS standard (code 94/580, E.coli O113:H10:K-), was used throughout the validation. This standard is identical to USP Reference Standard Endotoxin (EC6).
There are several possibilities to estimate the pyrogenic contamination of the preparations under test: 1) A quantitative estimation can be achieved by the construction of a dose-response curve for endotoxin standard (e.g. 5.0, 2.5, 1.0, 0.5 and 0.25 EU/ml) versus optical density (OD) value of the IL-6 ELISA. The contamination of the preparations is expressed in endotoxin–equivalent units. 2) A qualitative test can be achieved by the inclusion of an endotoxin threshold control (e.g. one fixed dilution of the standard curve) which allows for the classification in positive and negative samples (i.e. pyrogenic and

non-pyrogenic samples). 3) A qualitative test can also be achieved by inclusion of an appropriate positive product control.
A detailed description of analysis methods used during the validation of the test-method can be found in section 5 of the current BRD.

2.2 Rational for selected test components

Provide a detailed description and rationale, if appropriate, for the following aspects of the proposed test method:
2.2.1 Materials, equipment, and supplies needed.
The materials, equipment and supplies used for the WB/IL-6 test method are laboratory items, that will be already available in a routine QC laboratory. There is no need for sophisticated or dedicated laboratory equipment throughout the test.
For all steps in the procedure, excluding the ELISA procedure, the materials (e.g. tips, containers, solutions) which will be in close contact with samples and blood cells need to be sterile and pyrogen free. The materials, equipment and supplies are specified in the method protocol attached in Appendix A.1. It should be realized that equivalent devices may also be used and it is the user's responsibility to validate the equivalence.

Materials for part 1: Blood Incubation
Equipment
- Incubator (37°C, 5% CO_2 humidified air)
- Class 2 laminar flow sterile cabinet
- Centrifuge (suitable for 50 ml centrifuge tubes)
- Vortex

Consumables
- 30 ml syringe and a 40 mm, 21 gauge hypodermic needle.
- 50 ml centrifuge tube containing 10 IU heparine per 1 ml bloodsample.
- Serological pipettes (5, 10 and 25 ml)
- Polypropylene conical tubes
- Pipettes suitable for 50 or 100 µl
- 96-wells tissue culture plates
- WHO-LPS standard

Materials for part 2: ELISA procedure
Equipment
- Multichannel pipettor
- Microplate mixer
- Microplate washer
- Microplate reader capable of readings at the appropriate wavelength
- A software package facilitating data generation, analysis, reporting, and quality control

Consumables
- Graduated cylinder and plastic storage container for Buffered Wash Solution
- Tip-Tubs for reagent aspiration with Multichannel pipettor
- 96-wells microtiter plate
- Mouse monoclonal anti-IL-6 antibody from clone 16 (Novartis)

- Horseradish peroxydase conjugated sheep polyclonal anti-Il-6 antibody.
- Human Interleukin-6 standard
- Coating buffer, blocking buffer, dilution buffer, stopping solution and wash solution as detailed in the method protocol.

The IL-6 ELISA used is an in-house assay developed in the Novartis laboratory (participating in this study) and uses the WHO IL-6 international standard. Any commercially available IL-6 ELISA kit using the same standard or a standard calibrated versus it may be used (if validated for this in-vitro pyrogen test).
Including the appropriate positive and negative controls in each run ensures the reliability and accuracy of the WB/IL-6 test method. As a positive control a specified amount of the Endotoxin Standard is used. The assay should be considered acceptable only if the criteria described in the method protocol are met. Also the criteria for allowed variability of replicates within an assay have to be met. The IL-6 standard curve is an additional control of the performance of the assay.

2.2.2 Dose-selection procedures, including the need for any dose range-finding studies or acute toxicity data prior to conducting a study, if applicable.
For every kind of test compound the interference with human blood and the Il-6 ELISA is determined. For this purpose, a preliminary "dose finding" test is conducted to establish a suitable (interference free) dilution for every new test compound. For the validation study (as described in section 4 of this BRD), the tested products were diluted according to their known ELC, which was usually far beyond interfering concentrations. The ELCs of the tested products or drugs were calculated according to the European Pharmacopoeia. If no endotoxin limit is defined it can be estimated by dividing 350 EU by the maximum hourly dose (example: the maximum hourly dose is 100 mg/patient, then the estimated endotoxin limit is 350/100=3.5EU/mg).

2.2.3 Endpoint(s) measured.
The proposed test method is an *in vitro* model of the fever response mechanism. It determines the release of interleukin-6 (IL-6) by monocytoid cells present in human blood. IL-6 is released into the culture medium upon the interaction of pyrogens and specific receptors on the monocytoid cells. The measured endpoint IL-6 is one of the cytokines that trigger the fever response *in vivo*.

2.2.4 Duration of exposure.
The human whole blood is exposed to possible pyrogenic components in samples at 37°C for 16-24 hours in an atmosphere of 5% CO_2 in humidified air. The supernatant, containing endogenous pyrogens released by the cells, is subsequently assayed in the IL-6 ELISA.

2.2.5 Known limits of use.
The WB/IL-6 method described in the protocol in Appendix A is not a finalized test system for the testing of medicinal products. The method may be applied only to preparations that have been validated with this method, i.e. shown not to interfere with the blood and the IL-6 readout system at a specified dilution of the preparation. A

the blood and the IL-6 readout system at a specified dilution of the preparation. A paragraph describing the interference testing is included in the protocol (see Appendix A). However, at this moment there are no medicinal products known that can not be tested with the method.

2.2.6 Nature of the response assessed.
The proposed test method is an *in vitro* model of the fever response mechanism. Upon the interaction of exogenous pyrogens and specific receptors on the monocytoid cells endogenous pyrogens (e.g. interleukins, TNF-α and prostaglandins) are produced. In the body the fever response is triggered by these endogenous pyrogens. Immunoreactive IL-6, the measured endpoint for the current method, is one of these endogenous pyrogens.

2.2.7 *Appropriate vehicle, positive, and negative controls and the basis for their selection.*
Throughout the development and validation phase the test compounds are diluted in 0.9% (w/v) clinical saline. This 0.9% clinical saline is considered an appropriate vehicle as no interference with active substances of a drug is to be expected.
In addition the test includes several controls.
A negative control: 0.9% clinical saline (sodium chloride)
A positive control: WHO-LPS 94/580, 0.5 EU/ml in clinical saline.
A negative product control (NPC): clean, released batch for each drug.
A positive product control (PPC): test item spike with WHO-LPS (code 94/580) at 0.5 EU/ml.
The positive and negative controls are the same in every assay and are needed to establish the sensitivity of the test system. In addition, a product-based set of controls is used to reveal product-related interference.

2.2.8 Acceptable range of vehicle, positive and negative control responses and the basis for the acceptable ranges.

The standard curve of the endotoxin solution is to satisfy the criteria for linearity and range as described in the ICH guideline Q2B validation of analytical procedures: methodology, November 1996). For general applications the tests should satisfy additional criteria as specified in the WB/IL6 protocol (Appendix A to this BRD).

However, for the results described throughout this BRD the data were accepted and analyzed according to the procedures described in section 5.3 "Statistics". This procedure was chosen as it allowed for a harmonized analysis of comparable data which were obtained with different *in vitro* pyrogen tests (i.e. PBMC/IL-6, MM6/IL-6, WB/IL-1).

As regards the substances to be tested, for products with an established ELC, specified in EU/ml, the product is diluted to its maximum valid dilution (MVD). The negative product control should be negative at the MVD. The response to the positive product control should be between 50% and 200% of the response to the positive control, indicating a possible pyrogenicity can be detected using these conditions.

2.2.9 Nature of the data to be collected and the methods used for data collection.
The raw data collected are the read-outs (absorbance) of the IL-6 ELISA, measured by an automated laboratory ELISA-plate reader. The wavelength is dependent on the chromogenic substrate applied, but when using 3,3´,5,5´ -tetramethylbenzidine (TMB), the ELISA-plate is read at a wavelength of 450 nm. Bi-chromatic measurement with a reference wavelength of 540-590 nm is recommended.

2.2.10 Type of media in which data are stored.
Data are stored in electronic files (windows98 compatible software) and as hard copy.

2.2.11 Measures of variability.
As part of the development of the WB/IL-6 test method the intralaboratory repeatability was assessed by independent and identical replicated measurement of the different concentrations of WHO-LPS. Furthermore, the limit of detection and its dependence from known but uncontrollable variables such us operator and blood donor were investigated. These variables and the inherent variation of biological systems make up to the total variation of the method.

2.2.12 Statistical or non-statistical methods used to analyze the resulting data, including methods to analyze for a dose-response relationship. Justify and describe the method(s) employed.
All experiments are run with four replicates of the test compound with blood from one donor on one plate. A standard curve in quadruplicate, using the International Standard for Endotoxin (calibrated in EU) is included, ranging from 0.25 EU/ml up to 2.5 EU/ml. Outliers are rejected only after checking according to the Grubbs test, and applied to identify and eliminate aberrant data. Next, the negative and the respective positive control are compared to ensure a suitable limit of detection, which should be >0.25 EU/ml.

2.2.13 Decision criteria and the basis for the prediction model used to classify a test chemical (e.g., positive, negative, or equivocal), as appropriate.
A prediction model (PM) was developed in order to classify substances as “pyrogenic for humans” or “non-pyrogenic for humans”. To be able to define a dichotome result in the alternative pyrogen test, a threshold pyrogen value of 0.5 EU/ml was chosen. This threshold value was based on historical data with rabbits (described in section 4.1). The suitability of the PM was assessed by testing substances which were artificially contaminated with endotoxin (substances are described in section 3.2 and 3.3).
The statistical approach, including quality criteria, is detailed in section 5.3

2.2.14 Information and data that will be included in the study report and availability of standard forms for data collection and submission.
Raw data were collected using a standard form. These were submitted to the quality department of ECVAM

2.3 Basis for selection of this test method

Explain the basis for selection of the test method system. If an animal model is being used, this should include the rationale for selecting the species, strain or stock, sex, acceptable age range, diet, and other applicable parameters.

In view of the shortcomings of the rabbit pyrogen test and the BET, *in vitro* pyrogen tests that utilize the exquisite sensitivity to exogenous pyrogen of monocytoid cells have been proposed. In such tests, products are incubated with human cell and the conditioned media assayed for pyrogenic cytokines (Duff & Atkins, 1982; Dinarello et al., 1984; Poole et al., 1988; Poole et al, 1989; Hansen & Christensen, 1990; Taktak et al., 1991; Bleeker et al., 1994).

The human whole blood assay was developed as a real *in vitro* alternative to the rabbit pyrogen test. The basic idea was to mimic the fever reaction in humans. In general, the detection of exogenous pyrogens (e.g. endotoxin) by blood cells causes them to release endogenous pyrogens like IL-1β, IL-6 and TNFα. These cytokines affect the thermal regulation centre in the brain and increase the body temperature by changing its set point.

In the past, several test methods have been developed that use the sensitivity of human peripheral blood monocytes to exogenous pyrogens. In an attempt to increase the sensitivity of these tests the monocytes/leukocytes were isolated from whole blood. In addition, various cell lines, which retain monocytoid characteristics, including the capacity to synthesize and secrete pyrogenic cytokines, have been studied.

However, the isolation of monocytes/leukocytes from whole blood as well as the maintenance of a cell-line are labour–intensive and time–consuming, technically sophisticated and require expensive reagents. It is clear that using whole blood implies considerably simplified handling and that costs are limited.

An overview of relevant literature can be found in section 9 of this BRD. Interleukin IL-6 is chosen as the readout because IL-6, unlike IL-1 and TNF, is secreted entirely into the cell-conditioned medium in large quantities, thereby permitting its complete estimation.

2.4 Proprietary components

If the test method employs proprietary components, describe what procedures are used to ensure their integrity (in terms of reliability and accuracy) from "lot-to-lot" and over time. Also describe procedures that the user may employ to verify the integrity of the proprietary components.

S. Poole is named as an inventor in Patent Number US 6,696,261 B2 , Feb 24, 2004: 'Pyrogenicity test for use with automated immunoassay systems'.

T. Hartung and A. Wendel are named as inventors in Patent Number US 5,891,728 , Apr 6, 1999: 'Test for determining pyrogenic effect of a material'.

For clarification:

2.5 Replicates

Describe the basis for the number of replicate and repeat experiments; provide the rationale if experiments are not replicated or repeated.

All experiments are run with four replicates of the test compound on one plate. Outliers are rejected only after checking according to the Grubbs test ($p>0.05$). Four replicates is considered the minimal amount for the Grubbs test.

During a prevalidation phase, the intralaboratory reproducibility as well as the interlaboratory reproducibility of the WB/IL-6 test method was established by applying repeated experiments (see section 7). As the test method reliability (repeatability/reproducibility) was shown to be satisfactory, it was feasible to establish the accuracy using pharmaceutical substances (detailed in table 3.3.1) by one test performed by three participating laboratories (see section 6).

2.6 Modifications applied after validation

Discuss the basis for any modifications to the proposed test method protocol that were made based on results from validation studies.
The test can easily be adjusted to a quantitative assay as described in the method protocol. However, the assay has now been validated as a qualitative assay by means of the PM.

2.7 Differences with similar test methods

If applicable, discuss any differences between the protocol for the proposed test method and that for a comparable validated test method with established performance standards.
Not applicable.

3 Substances Used for Validation

3.1 Selection of substances used

Describe the rationale for the chemicals or products selected for use in the validation process. Include information on the suitability of the substances selected for testing, indicating any chemicals that were found to be unsuitable.

Selected test items were medicinal products available on the market. Released clinical batches were considered clean, i.e. containing no detectable pyrogens. To test the specificity, sensitivity and the reproducibility of the proposed test method, the products were spiked with pyrogen. For the present studies endotoxin (LPS) was selected as the model pyrogen, since it is well defined, standardized and readily available.

For the sensitivity and specificity the test items were assessed at their MVD. The MVD is the quotient of the ELC and the detection limit. The European Pharmacopoeia prescribes for various types of parenterals the amount of endotoxin that is maximally allowed in a medicinal product, i.e. the ELC, taking into consideration the dose, the route of administration and the dosing regimen of the product.

The aim of the study was to discriminate between negative and positive samples. Based on the selected pyrogen threshold value (see 4.1) and the PM applied, positive samples were defined as containing 0.5 EU/ml or more. Hence, to determine the MVD, the value of 0.5 EU/ml was defined as the detection limit.

Test items were assessed as such (negative product control), spiked with endotoxin at 0.5 IU/ml (positive product control) and after spiking with endotoxin at 5 levels (blinded samples). In addition a negative control (saline) and positive control (0.5 EU/ml in saline) were included to establish assay validity.

For reproducibility, the test items were tested, at a predefined dilution above the MVD, independently in 3 different laboratories, 3 times each. Based on the selected pyrogen threshold value (see 4.1) and the PM applied, positive samples were defined as containing 0.5 EU/ml or more. The test items were tested after spiking with endotoxin at four levels. For no other reasons but practical ones, i.e. availability of test materials, different test items were selected for this part of the validation study.

It was determined earlier whether candidate test items interfered with the outcome of the proposed test method. Interference was considered when the response of endotoxin in the diluted test item was below 50% or above 200% of the response of endotoxin in saline (spike-recovery). It was shown that none of the test items interfered with the assay at the selected dilutions (data not shown).

3.2 Number of substances

Discuss the rationale for the number of substances that were tested.

A total of 13 test items were selected for the validation study (see 3.3): 10 test items for determining sensitivity and specificity (table 3.3.1), 3 different test items for determining reproducibility (table 3.3.2). Test items and their spikes were appropriately blinded by ECVAM before distribution to the participating testing facilities.

For sensitivity and specificity, each test item was tested after spiking at its individual MVD. Hence they each came with their own specific set of 5 endotoxin spike solutions: 0.0, 0.25, 0.5, 0.5 and 1.0 EU/ml. Simple logistics limited the amount of test items for this part of the validation study to 10. Since test items were assessed with 5 different endotoxin levels at 3 independent test facilities, this yielded a total of 150 data points, biometrically considered to be sufficient for further analysis.

For reproducibility each test item was spiked at 4 different levels (0.0, 0.0, 0.5 and 1.0 EU/ml) and tested at specified dilutions, 3 times at 3 laboratories.

3.3 Description of substances used

Table 3.3.1: Test items (parenteral drugs) used for determining sensitivity and specificity

Drug	code	Source	Agent	Indication	MVD (-fold)
Glucose 5% (w/v)	GL	Eifel	Glucose	nutrition	70
Ethanol 13% (w/w)	ET	B.Braun	Ethanol	diluent	35
MCP®	ME	Hexal	Metoclopramid	antiemetic	350
Orasthin®	OR	Aventis	Oxytocin	initiation of delivery	700
Binotal®	BI	Aventis	Ampicillin	antibiotic	140
Fenistil®	FE	Novartis	Dimetindenmaleat	antiallergic	175
Sostril®	SO	GlaxoSmithKline	Ranitidine	antiacidic	140
Beloc®	BE	Astra Zeneca	Metoprolol tartrate	heart dysfunction	140
Drug A*	LO	-	0.9% NaCl	-	35
Drug B*	MO	-	0.9% NaCl	-	70

*Drugs A and B were included as saline controls using notional ELCs.
Negative control: 0.9% clinical stock saline solution.
Positive control: WHO-LPS 94/580 (*E. coli* 0113:H10:K-), 0.5 EU/ml in clinical stock saline.

Table 3.3.2: Test items (parenteral drugs) used for determining reproducibility.

Drug	Source	Agent	Indication
Gelafundin®	Braun melsungen	Gelatin	Transfusion
Jonosteril ®	Fresenius	Electrolytes	Infusion
Haemate ®	Aventis	Factor VIII	Hemophilia

Negative control: 0.9% clinical stock saline solution.
Positive control: WHO-LPS 94/580 (*E. coli* 0113:H10:K-), 0.5 EU/ml in clinical stock saline.

3.4 Sample coding procedure

Describe the coding procedures used in the validation studies.

All test items are registered medicinal products and were obtained from a pharmaceutical supplier. Test items and endotoxin spiking samples were prepared, blinded where appropriate and coded under GLP by personnel from ECVAM, Italy. These were then taken over by the Paul-Ehrlich Institute, Germany, for allocation and shipment to each of the appropriate test facilities participating in the study.
For the sensitivity and specificity part of this study, test items and their respective spikes (5 per test item) were all blinded. For reproducibility testing, only the spikes (4) were blinded, the test items were not.

3.5 Recommended reference chemicals

For proposed test methods that are mechanistically and functionally similar to a validated test method with established performance standards, discuss the extent to which the recommended reference chemicals were tested in the proposed test method. In situations where a listed reference chemical was unavailable, the criteria used to select a replacement chemical should be described. To the extent possible, when compared to the original reference chemical, the replacement chemical should be from the same chemical/product class and produce similar effects in the in vivo reference test method. In addition, if applicable, the replacement chemical should have been tested in the mechanistically and functionally similar validated test method. If applicable, the rationale for adding additional chemicals and the adequacy of data from the in vivo reference test method or the species of interest should be provided.
The reference pyrogen material used was the international endotoxin standard WHO-LPS 94/580 (*E. coli* 0113:H10:K-). Where appropriate, the material was diluted in clinical saline solution (0.9%(w/v) sodium chloride). The saline was also used as negative control (blank).

4 *In vivo* Reference Data on Accuracy

4.1 Test protocol *in vivo* reference test method.

Provide a clear description of the protocol(s) used to generate data from the in vivo reference test method. If a specific guideline has been followed, it should be provided. Any deviations should be indicated, including the rationale for the deviation.
For ethical reasons, no rabbit pyrogen tests were performed for this study. However, Dr. U. Lüderitz-Püchel, Paul-Ehrlich Institute, Germany, kindly provided historical data, accumulated over several years, from 171 rabbits (Chinchilla Bastards). The respective Pharmacopoeia's do not prescribe a rabbit strain for the *in vivo* pyrogen test, but Chinchilla rabbits are reported as a relatively sensitive strain for pyrogen testing.

The rabbits were injected with endotoxin and their rise in body temperature over the next 180 minutes was recorded (figure 4.1.1). From these data it was established that 50% of the rabbits got fever when treated with endotoxin at 5 EU/kg (Hoffmann et al, 2005a). Fever in rabbits is defined as a rise in body temperature over 0.55°C. On the basis of these historical animal data and corrected for the maximal volume allowed in rabbits, i.e. 10 ml/kg per animal, a pyrogen threshold value of 0.5 EU/ml was defined for the PM in the proposed test method.

4.2 Accuracy

Provide the in vivo reference test method data used to assess the accuracy of the proposed test method. Individual human and/or animal reference test data, if available, should be provided. Provide the source of the reference data, including the literature citation for published data, or the laboratory study director and year generated for unpublished data.
As mentioned, no animal studies were done for ethical reasons. However, a theoretical assumption on specificity and sensitivity can be made (Hoffmann et al, 2005a) Taking the prevalence of the different final concentrations of LPS in the 5 spikes into account (1.0 EU/ml: 20%; 0.5 EU/ml: 40%; 0.25 EU/ml: 20% and 0.0 EU/ml: 20%) and calculating probabilities of misclassification using the chosen study design and defined threshold of pyrogenicity, i.e. 0.5 EU/ml, the theoretical sensitivity of the rabbit pyrogen test is 57.9% and the theoretical specificity of the rabbit pyrogen test is 88.3%.

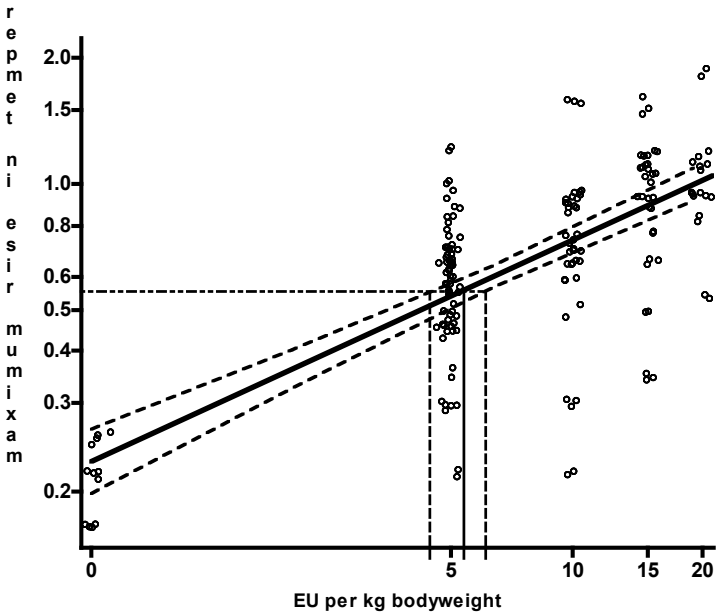


Figure 4.1.1 Dose-temperature of standard endotoxin applied to Chinchilla Bastards (n=171). Rabbits were treated with 1 ml saline containing 0, 5, 10, 15 and 20 EU of *E. coli* LPS (WHO-LPS 94/580 (E.coli O113:H10:K)) and their body temperature was measured over 180 min. Linear regression analysis was performed after logarithmic transformation of the data. Data are shown as dots to which a jitter-effect was applied in order to be able to distinguish congruent data. The full line depicts the linear regression whereas the dashed lines represent the 95%-confidence bounds. Furthermore, a horizontal line for a 0.55°C raise of temperature is added which is often defined as the rabbit threshold for fever. At the interception point of this line and the regression line 50% of the rabbits are to be expected to develop fever.

4.3 Original records

If not included in the submission, indicate if original records are available for the in vivo reference test method data.

The recognition of pyrogenic substances as bacterial by-products and the identification of a variety of pyrogenic agents enabled the development of a proper test to demonstrate non-pyrogenicity of the pharmaceutical product. As early as the 1920s, studies were done to select the most appropriate animal model. Results indicated that most mammals had a pyrogenic response, but only a few, including rabbits, dogs, cats, monkeys and horses showed a response similar to that in humans. For practical reasons, other species but rabbits and dogs were considered not practical. In 1942, Co Tui % Schrift described that rabbits are less thermo-stable as compared to dogs. Hence, rabbits are more suited for the purpose of testing for the absence of pyrogens, since a negative result is more significant.

4.4 Quality of data

Indicate the quality of the in vivo reference test method data, including the extent of GLP compliance and any use of coded chemicals.

Not applicable.

4.5 Toxicology

Discuss the availability and use of relevant toxicity information from the species of interest (e.g., human studies and reported toxicity from accidental or occupational exposure for human health-related toxicity testing).

Over time, a number of studies were done to correlate the rabbit test to pyrogenic reactions in humans. A conclusive study by Greisman and Hornick, published in 1969, who compared three purified endotoxin preparations (*Salmonella typhosa*, *E. Coli* and *Pseudomonas*) in New Zealand rabbits and in male volunteers, showed that the induction of a threshold pyrogenic response, on a weight basis, was similar to rabbit and man. At higher doses, rabbits respond less severe as compared to man.

4.6 Background on assay performance

Discuss what is known or not known about the accuracy and reliability of the in vivo reference test method.

As mentioned, no animal studies were done for ethical reasons. However, a theoretical assumption on specificity and sensitivity can be made (Hoffmann et al 2005a). Taking the prevalence of the different final concentrations of LPS in the 5 spikes into account (1.0 EU/ml: 20%; 0.5 EU/ml: 40%; 0.25 EU/ml: 20% and 0.0 EU/ml: 20%) and calculating probabilities of misclassification using the chosen study design and defined threshold of pyrogenicity, i.e. 0.5 EU/ml, the theoretical sensitivity of the rabbit pyrogen test is 57.9% and the theoretical specificity of the rabbit pyrogen test is 88.3%.

5 Test Method Data and Results

5.1 Test method protocol

Describe the proposed test method protocol used to generate each submitted set of data. Any differences from the proposed test method protocol should be described, and a rationale or explanation for the difference provided. Any protocol modifications made during the development process and their impact should be clearly stated for each data set.

The method protocol for the WB/IL-6 test is provided in the Appendix A of this BRD. It includes the precise step-by-step description of the test method, including the listing of all the necessary reagents and laboratory procedures for generating data. For two steps during validation a part of the protocol was adapted to contain a detailed description of the dilution of the samples and the spiking with WHO-LPS. The relevant part is detailed in this section. The validity criteria and the detailed statistical analysis described in section 5.3 of this BRD were applied to analyse the data produced during validation.
To assess the reliability of the test method a series of experiments were conducted in the DL. As a start, only blanks (saline, 0 EU/ml) and spikes of WHO-LPS in saline were tested. These experiments are summarised in table 5.1.1.

Table 5.1.1 : summary of experiments with WHO-LPS in saline

Experiment	Spikes (EU/ml) in saline	n (per spike)	Repetitions of experiment	N
1A	0; 0.5	20	1	40
1B	0; 0.063; 0.125; 0.25; 0;5	10	1	50
2A	0; 0.25; 0.5	8	3	72
2B	0; 0.5	5	8	80

The collected data were used to answer questions regarding the nature of the distribution, the variance and its behaviour over the range of response in replicated measurements under identical conditions. In addition, intralaboratory reproducibility was assessed by the maintenance of minimum criteria, e.g. a detection limit below an a-priori chosen positive control or a dose-dependent standard curve (table 5.1.1, experiment 1b). With the data of this experiment an assessment of the limit of detection of the test can be done by calculating the smallest spike, which can be discriminated from the blank. Intralaboratory reproducibility was assessed by the maintenance of minimum criteria, e.g. a detection limit below an a-priori chosen positive control of a dose dependent standard curve.

Next, the WB/IL-6 method was transferred from the DL to two other laboratories (denoted as naive laboratory 1 [NL1] and naive laboratory 2 [NL2]). All three laboratories performed a large-scale dose response experiment. For this study 6 or 7 concentrations were tested in a dose response curve (typically 0; 0.125; 0.25; 0.5; 1; 2 EU/ml, at least 8 replicates) and all laboratories had to meet the validity criteria as laid down in the protocol before the studies with medicinal substances were conducted.

The actual **intra- and interlaboratory reproducibility** was assessed by testing 3 different medicinal substances Gelafundin, Jonosteril and Haemate (described in table 3.3.2, section 3.3.). Test items and their spikes were appropriately blinded. Test items were tested, at a predefined dilution above the MVD, independently in 3 laboratories, 3 times each. Test items were tested after spiking with WHO-LPS at four different levels, the spikes were blinded and coded by QA ECVAM. In addition a negative control (saline) and positive control (0.5 EU/ml) in saline were included to establish assay validity.
Although this part of the study was designed for assessment of reproducibility, a preliminary estimate of the accuracy could be derived from the data. Applying the PM to the results and evaluating the concordance in a two-by-two contingency table assessed accuracy.

To assess **accuracy** of the proposed test method 10 substances (listed in table 3.3.1), were spiked with five different concentrations of the WHO-LPS (one of which is negative). To permit the application of the chosen PM, each drug was diluted to its individual MVD, which has been calculated using the ELC to that drug (listed in table 3.3.1.) Each substance had their own specific set of endotoxin spike solution, ensuring that after spiking the undiluted drug, it contained 0.0; 0.25; 0.5; 0.5 and 1.0 EU/ml respectively when tested at its MVD. A detailed description of the sample preparation was supplied to the three independent laboratories (shown in table 5.1.2). To put more weight to the result of this part of the validation, the spikes were blinded and coded by QA ECVAM. The raw data of the WB/IL-6 assay are shown in paragraph 5.2. Accuracy was assessed by applying the PM to the results and evaluating the concordance in a two by two table. As intralaboratory reproducibility was (successfully) shown in previous experiments, only interlaboratory reproducibility was assessed in this phase.

Table 5.1.2: Sample preparation for the testing of 10 substances spiked with 5 different concentrations of WHO-LPS.

<table>
<tr><td colspan="3">unblinded</td><td colspan="5">blinded</td></tr>
<tr><td colspan="3">dilution of drug up to MVD
↓</td><td colspan="5">spiking of undiluted drug: 0.5 ml each</td></tr>
<tr><td>diluted drug</td><td>NPC</td><td>PPC</td><td rowspan="2">+ 23.3 µl of Spike 1</td><td rowspan="2">+ 23.3 µl of Spike 2</td><td rowspan="2">+ 23.3 µl of Spike 3</td><td rowspan="2">+ 23.3 µl of Spike 4</td><td rowspan="2">+ 23.3 µl of Spike 5</td></tr>
<tr><td rowspan="2">0.5 ml</td><td>+ 25 µl saline</td><td>+ 25 µl PPC-LPS-spike *</td></tr>
<tr><td></td><td>(final conc. = 50 pg/ml)</td><td colspan="5">dilution to MVD
↓</td></tr>
<tr><td></td><td>test</td><td>test</td><td>test</td><td>test</td><td>test</td><td>test</td><td>test</td></tr>
</table>

** PPC-LPS-spike contains 1050 pg/ml = 21fold 50 pg/ml*

NPC = Negative Product Control, **PPC** = Positive Product Control, **MVD** = Maximal Valid Dilution

5.2 Accuracy and reliability

Provide all data obtained to evaluate the accuracy and reliability of the proposed test method. This should include copies of original data from individual animals and/or individual samples, as well as derived data. The laboratory's summary judgment regarding the outcome of each test should be provided. The submission should include data (and explanations) from all studies, whether successful or not.
See figures 5.2.1, 5.2.2, 5.2.3, 5.2.4, 5.2.5 (A, B and C), 5.2.6 and 5.2.7 (A and B).

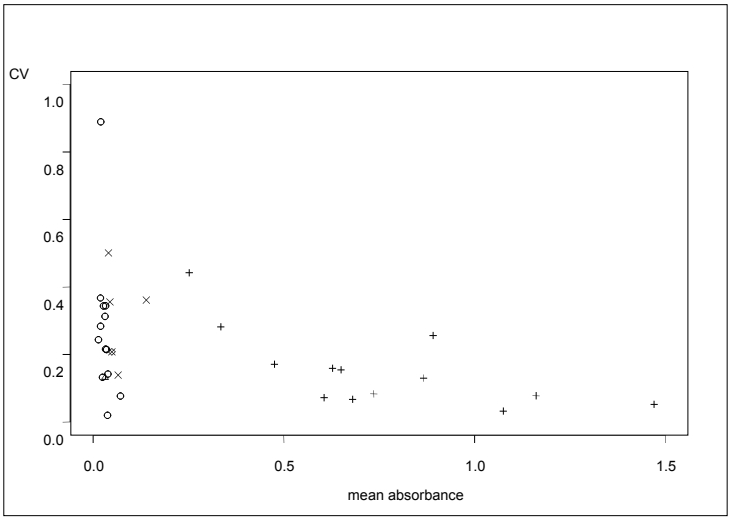


Figure. 5.2.1: Coefficient of variation (CV) of WHO-LPS spikes relative to the mean absorbance (readout of the IL-6 ELISA).

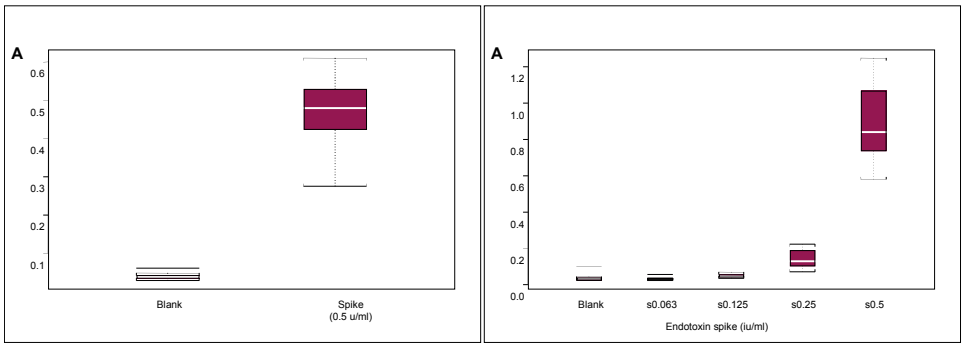


Figure. 5.2.2: Boxplots with absorbance (A) values of 20- replicates (left) or 10 replicates (right) of WHO-LPS spikes in saline at various concentrations (readout of the IL-6 ELISA).

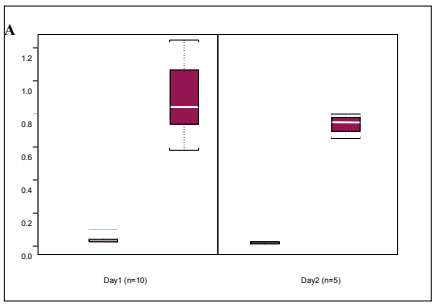


Figure. 5.2.3: Boxplot of absorbance (A) values of the response of two different blood donations from one single volunteer with WHO-LPS (IU/ml) in saline at 0.0 IU/ml and 0.5 IU/ml (readout of the IL-6 ELISA).

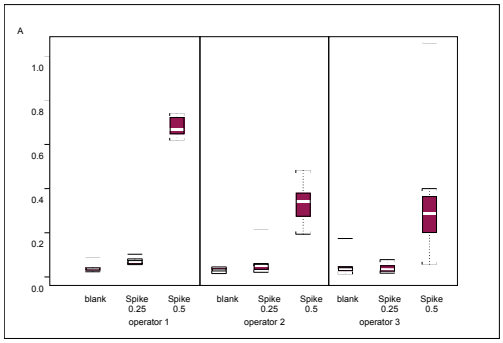


Figure. 5.2.4: Boxplot of absorbance (A) values of WHO-LPS (IU/ml) in saline at 0.0 (Blank) , 0.25 IU/ml or 0.5 IU/ml. with 3 different operators (readout of the IL-6 ELISA).

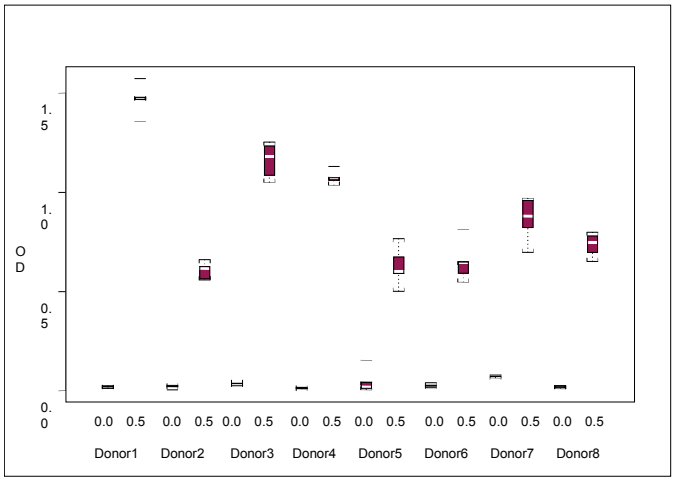


Figure. 5.2.5: Boxplot of absorbance (A) values of the response of 8 individual donors to WHO-LPS (IU/ml) in saline at 0.0 IU/ml and 0.5 IU/ml (readout of the IL-6 ELISA).

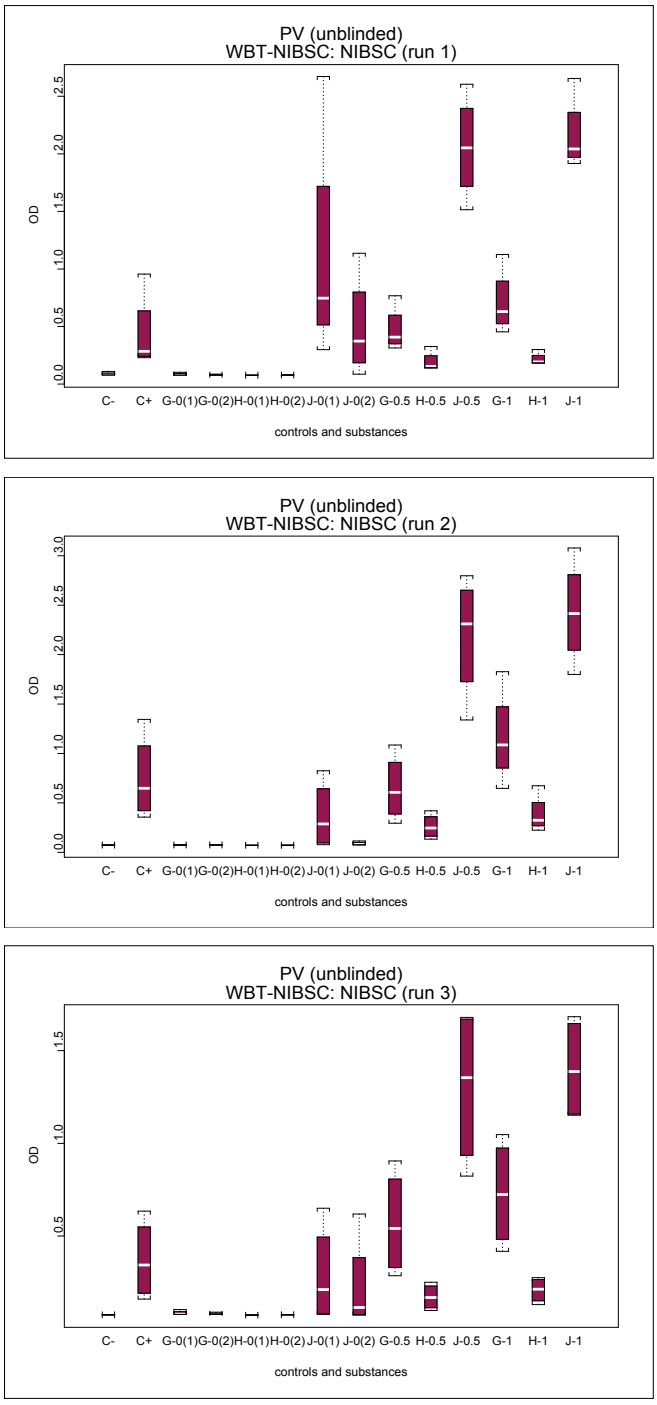


Figure. 5.2.6 A: Three different drugs were spiked (blinded) with WHO-LPS at 0.0, 0.0, 0.5 and 1.0 IU/ml, respectively. Experiment was run 3 time independently at the NIBSC laboratory (readout of the IL-1 ELISA).
G = Gelafundin; J = Jonestreril; H = Heamate.
C- = negative control (saline); C+ = positive control (0.5 IU/ml in saline).

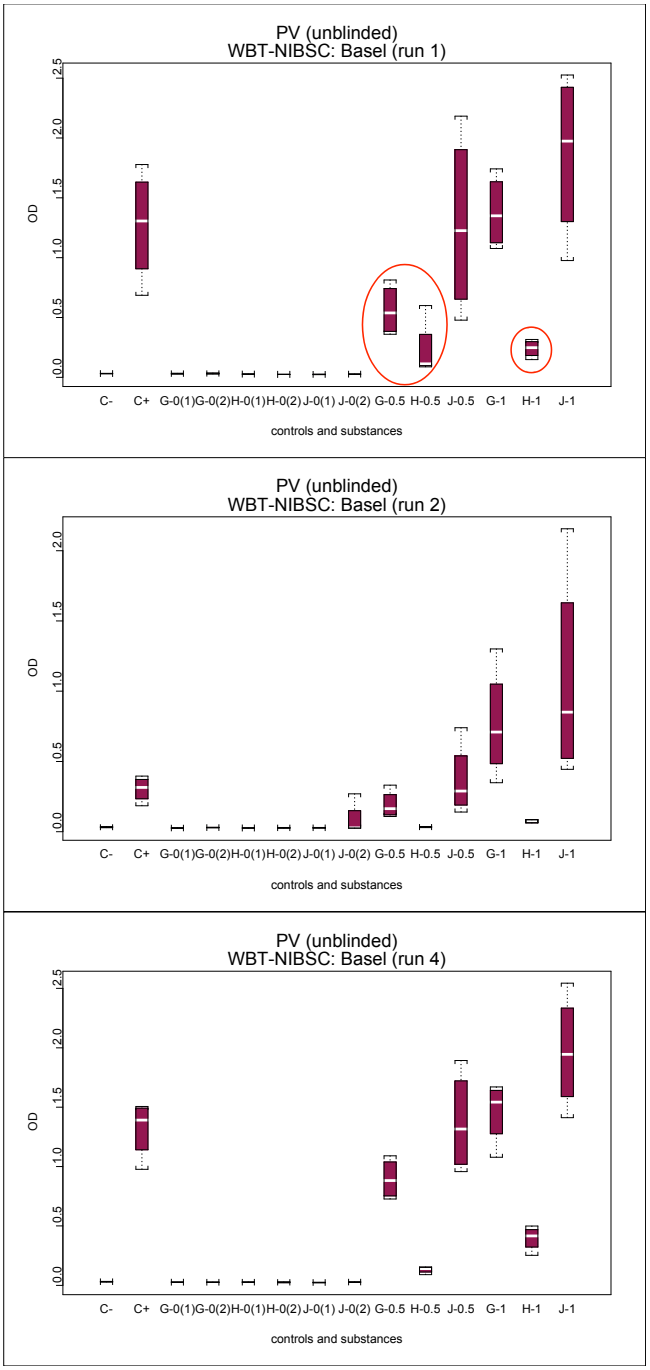


Figure. 5.2.6 B: Three different drugs were spiked (blinded) with WHO-LPS at 0.0, 0.0, 0.5 and 1.0 IU/ml, respectively. Experiment was run 4 times independently at the Basel laboratory (readout of the IL-6 ELISA). The third run (results not shown) was declared invalid for technical reasons.
G = Gelafundin; J = Jonestreril; H = Heamate.
C- = negative control (saline); C+ = positive control (0.5 IU/ml in saline).

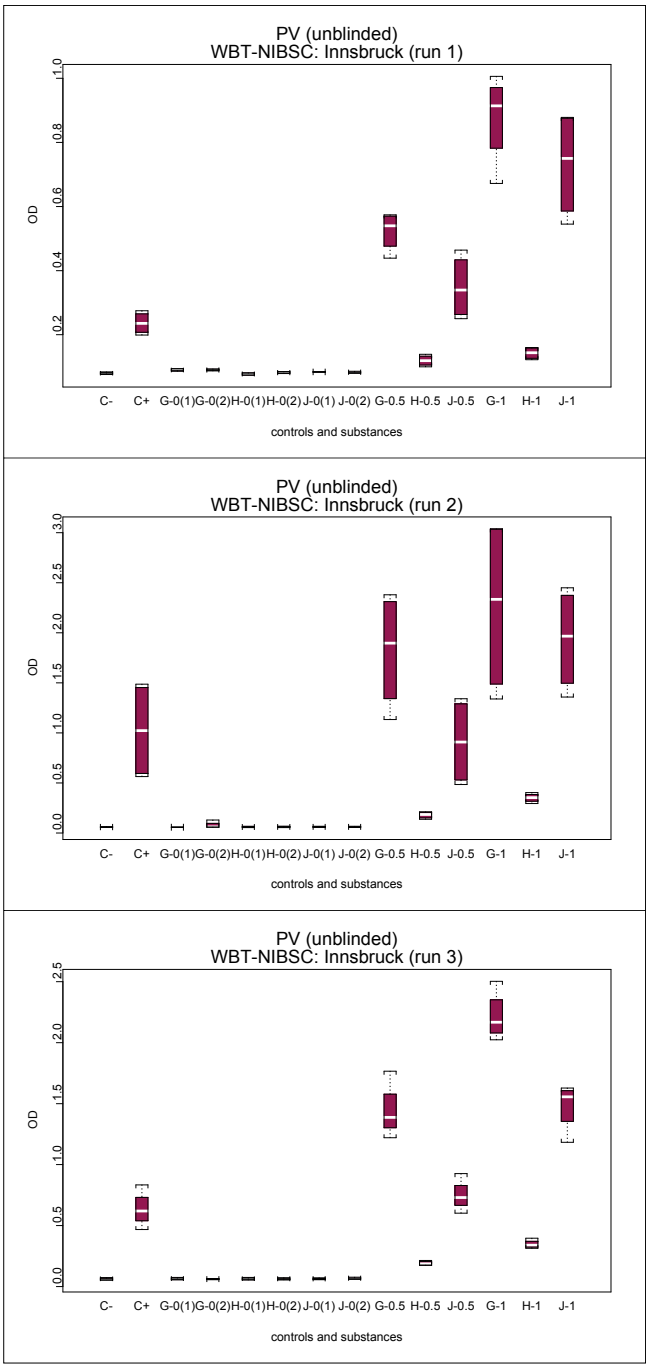


Figure. 5.2.6 C: Three different drugs were spiked (blinded) with WHO-LPS at 0.0, 0.0, 0.5 and 1.0 IU/ml, respectively. Experiment was run 3 time independently at the Innsbruck laboratory (readout of the IL-6 ELISA).
G = Gelafundin; J = Jonestreril; H = Heamate.
C- = negative control (saline); C+ = positive control (0.5 IU/ml in saline).

BRD: WB/IL-6 March, 2006

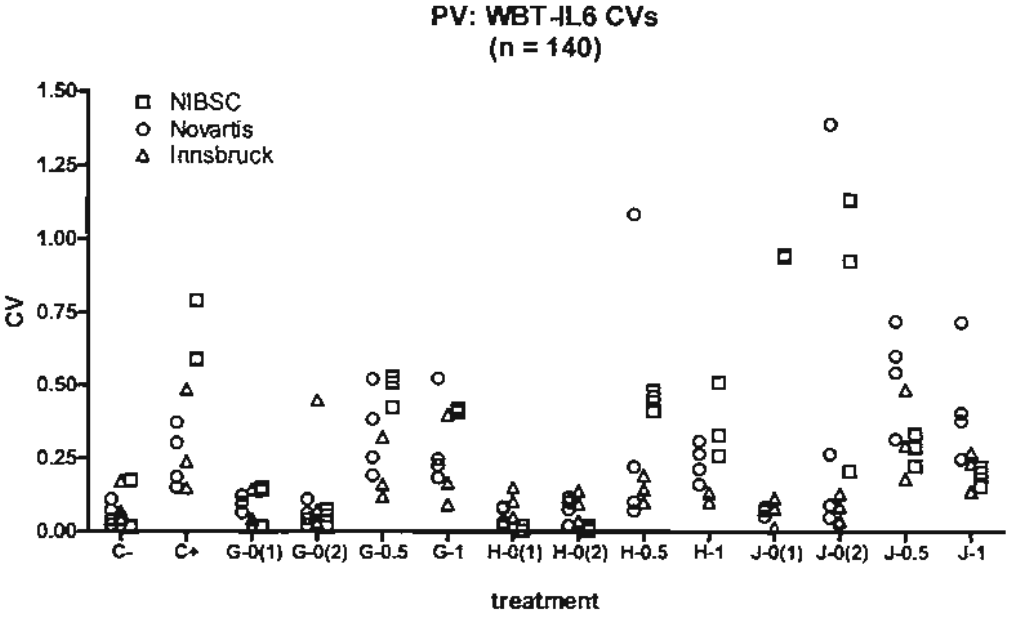


Figure. 5.2.7: Coefficient of variation (CV) of different WHO-LPS spikes (0.0, 0.0, 0.5 and 1.0 IU/ml, respectively).from the experiments as shown in fig. 5.2.6 A-C.
G = Gelafundin; J = Jonestreril; H = Heamate.
NC = negative controle (saline); PC is positive conrole (0.5 IU/ml in saline).

5.3 Statistics

Describe the statistical approach used to evaluate the data resulting from studies conducted with the proposed test method.

A generally applicable analytical procedure was employed. This procedure includes a universal PM as well as quality criteria. First, a two-step procedure consisting of a variance-criterion and an outlier-test was applied. For this, the Dixon's test (Barnett and Lewis, 1984), which is USP approved, was chosen with the significance level of α=0.01 and applied to identify and eliminate aberrant data.

Next, the negative and the respective positive control are compared to ensure a suitable limit of detection. For this, a one-sided t-test with a significance level of α=0.01 is applied to the ln-transformed data to ensure that the response to the positive control is significantly larger than that of the respective negative control.

Finally, the samples are classified as either negative or positive by the outcome of a one-sided version of the t-test, which is based on the assigned pyrogen threshold value. The final results will be given in 2 x 2 contingency tables (table 5.3.1). These tables allow for estimation of accuracy (sensitivity and specificity) and reproducibility of the proposed test method.

Table 5.3.1: 2x2 contingency table.

		pre-defined class ("truth")		Σ
		1	***0***	
Classification by test system and PM	***1***	***a***	***b***	$a+b = n_{.1}$
	0	***c***	***d***	$c+d = n_{.0}$
Σ		$a+c = n_{1.}$	$b+d = n_{0.}$	***n***

Accuracy:
The most important statistical tool to determine accuracy (specificity and sensitivity) is the PM (Hothorn, 1995). In general, it is a statistical model, which classifies a given drug by an objective diagnostic or deciding rule. The objective of a dichotomous result requires a clear cut PM, which assigns a drug in one of the two classes "pyrogenic for humans" and "non-pyrogenic for humans". Since a threshold pyrogen value will be used, a one-sided test is appropriate for the task. Because the data are normalised by a ln-transformation, a t-test is chosen. Although the variances over the range of concentration converge by the transformation, the assumptions of equal variances do generally not hold true, because it depends on additional covariates. Therefore, the one sided Welch-t-test (Snedecor and Cochran, 1989) is applied. Due to the safety aspect of the basic problem, the hypotheses of the test are

$$H_0 : \mu_{S_i j} > \mu_{S+} \quad vs \quad H_1 : \mu_{S_i j} < \mu_{S+},$$

where $\mu_{...}$ denotes the parameter of location of the respective ln-transformed distribution. This approach controls the probability of false positive outcomes directly by means of its significance level α, which is chosen as 0.01, because is assumes hazard, respectively pyrogenicity, of the tested drug in H_0, and assures safety, i.e. non-pyrogenicity. The test statistic is

$$T_{S_i j} = \frac{\bar{x}_{S+} - \bar{x}_{S_i j}}{\sqrt{\frac{s^2_{S+}}{n_{S+}} + \frac{s^2_{S_i j}}{n_{S_i j}}}}.$$

The PM is built by means of the outcome of the test. Let 0 denote safety and 1 denote hazard. The classification of S_i-j is then determined by

$$S_i j = 0, \text{ if } \; T_{S_i j} > t_{0.99; n_{S+} + n_{S_i j} - 2},$$

$$S_ij = 1, else,$$

where $t_{0.99;n_{S+}+n_{S_ij}-2}$ the 0.99-quantile of the t-distribution with $n_{S+} + n_{S_ij} - 2$ degrees of freedom. The number of replicates for every control and sample, i.e. $n_{...}$, was harmonised to be four. Due to the possibility of removing one observation by the outlier test, the number of replicates could be reduced to three. The classification of a version of a drug is regarded as an independent decision. Therefore, the niveau α is local.

Finally, the classifications of the drugs will be summarised in 2x2 contingency table (table 3). From these tables, estimates of the sensitivity (S_E), i.e. the probability of correctly classified positive drugs and specificity (S_P), i.e. the probability of correctly classified negative drugs, will be obtained by the respective proportions. Where

$$S_E = a / (a + c) * 100\%$$
and
$$S_P = d / (b + d) * 100\%.$$

Furthermore, these estimates will be accompanied by confidence intervals, which will be calculated by the Pearson-Clopper method (Clooper & Pearson, 1934). For example, let $\hat{p}_{SE}$ denote the proportion, namely the sensitivity, under investigation. Then the confidence interval to a niveau α is calculated as

$$\left[p_{SE}^{L} = \frac{\mathrm{a}F_{2\mathrm{a};2(n_{1\cdot}-\mathrm{a}+1);\frac{\alpha}{2}}}{n_{1\cdot} - \mathrm{a} + 1 + \mathrm{a}F_{2\mathrm{a};2(n_{1\cdot}-\mathrm{a}+1);\frac{\alpha}{2}}}; p_{SE}^{U} = \frac{(\mathrm{a}+1)F_{2(\mathrm{a}+1);2(n_{1\cdot}-\mathrm{a});1-\frac{\alpha}{2}}}{n_{1\cdot} - \mathrm{a} + (\mathrm{a}+1)F_{2(\mathrm{a}+1);2(n_{1\cdot}-\mathrm{a});1-\frac{\alpha}{2}}} \right],$$

where $F_{...}$ denotes the respective quantile of the F-distribution and $n_{1\cdot}$ is the sample size of the positive drugs and a the number of correctly classified drugs.

By contaminating the drugs artificially and by defining a threshold value, which is assumed to be appropriate, the class of a drug is determined beforehand. The versions of drugs, which are effectively contaminated, but below the threshold dose, are considered to be negative, respectively safe, because their contamination is not crucial for humans in terms of ELC.

Reproducibility:

The analysis of the intra- and interlaboratory reproducibility was assessed from the three identical and independent runs conducted in each of 3 laboratories. The comparison of the three runs was carried out blindly such that the testing facility did not know the true classification of the sample, either pyrogenic or non-pyrogenic. By this procedure only the randomness and reproducibility of the methods was assessed and not systematic errors, which may have arisen from other sources, e.g. logistics or preparation of the samples. This means that if a sample was (mis)classified in all three runs the result is reproducible regardless of the (mis)classification of the sample. Therefore, a measure of

similarity, i.e. complete simple matching with equal weights, was preferred to the coefficient of correlation for 2x2 contingency tables.
The study was designed as follows: each laboratory had to conduct three independent runs with the same 12 samples (3 test items with 4 blinded spikes each) and two controls, i.e. saline as a negative control (C-) and a 0.5 EU/ml LPS-spike in saline as a positive controls (C+). The samples were derived from the three substances Gelafundine, Haemate and Jonosteril. Per run, each substance was blindly spiked twice with saline, once with 0.5 EU/ml LPS and once with 1 EU/ml LPS, which resulted in a balanced design with regard to positive and negative samples, i.e. samples expected to be pyrogenic and non-pyrogenic, respectively.

The three independent runs per testing facility provide the information on which the assessment of the intralaboratory reproducibility is based. The combined results of the three runs per testing facility were used to determine interlaboratory reproducibility. The correlation of the prediction (in terms if the Bravais-Pearson coefficient of correlation) between all runs is calculated, independent of whether that classification is true or false. A BP-correlation of 1 is calculated, if two runs gave exactly the same predictions for the twelve substances. If one run gives adverse classifications for all substances than the other, the correlation is –1. As these calculations do not need information of the true status of a sample, they were carried out blinded.

5.4 Tabulated results

Provide a summary, in graphic or tabular form, of the results.
See tables 5.4.1 and 5.4.2.

Table 5.4.1: Results of testing 3 substances 3 times by 3 laboratories. Classifications after applying the prediction model (compare to fig. 5.2.5)

Sample	DL (NIBSC)			NL 1 (Basel)			NL 2 (Innsbruck)		
	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3
G-0 (1)	0	0	0	0	0	0	0	0	0
G-0 (2)	0	0	0	0	0	0	0	0	0
H-0 (1)	0	0	0	0	0	0	0	0	0
H-0 (2)	0	0	0	0	0	0	0	0	0
J-0 (1)	1	0	1	0	0	0	0	0	0
J-0 (2)	1	0	1	0	0	0	0	0	0
G - 0.5	1	1	1	0	1	1	1	1	1
H - 0.5	1	0	1	0	0	0	0	0	0
J - 0.5	1	1	1	1	1	1	1	1	1
G - 1	1	1	1	1	1	1	1	1	1
H - 1	1	1	1	0	0	0	0	0	0
J - 1	1	1	1	1	1	1	1	1	1

"0"denotes "non-pyrogenic"; "1" denotes "pyrogenic".

Table 5.4.2:

Results of the validation study of 10 drugs, spiked with WHO-LPS at 0.0, 0.25, 0.5, 0.5 and 1.0 IU/ml, respectively and tested in 3 different laboratories. Samples and spikes were blinded. Classifications after applying the prediction model (compare to fig. 5.2.7).

drug (code)	spike		results		
	EU/ml	"truth"	PEI	Basel	Innsbruck
Beloc (BE)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**1**	**0**	**0**
	0.50	**1**	**1**	**1**	**0**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
Binotal (BI)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**0**
	0.50	**1**	**1**	**0**	**0**
	0.50	**1**	**1**	**0**	**0**
	1.00	**1**	**1**	**1**	**1**
Ethanol 13% (ET)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**0**
	0.50	**1**	**1**	**1**	**1**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
Fenistil (FE)	**0.00**	**0**	**0**	**0**	**0**

drug (code)	spike EU/ml	"truth"	results PEI	Basel	Innsbruck
	0.25	0	0	0	0
	0.50	1	1	0	1
	0.50	1	1	0	1
	1.00	1	1	1	1
Glucose 5% (GL)	0.00	0	0	0	0
	0.25	0	0	0	0
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
"Drug A" 0.9% NaCl (LO)	0.00	0	0	0	0
	0.25	0	0	0	0
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
MCP (ME)	0.00	0	0	0	0
	0.25	0	NA	0	0
	0.50	1	1	1	1
	0.50	1	NA	1	1
	1.00	1	1	1	1
"Drug B" 0.9% NaCl (MO)	0.00	0	0	0	0
	0.25	0	0	0	0
	0.50	1	1	0	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Orasthin (OR)	0.00	0	0	0	0
	0.25	0	1	0	0
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Sostril (SO)	0.00	0	0	0	0
	0.25	0	0	0	0
	0.50	1	1	1	0
	0.50	1	1	1	0
	1.00	1	1	1	1

"0"denotes "non-pyrogenic"; "1" denotes "pyrogenic"; NA is not assessed.

5.5 Coding of data

For each set of data, indicate whether coded chemicals were tested, whether experiments were conducted without knowledge of the chemicals being tested, and the extent to which experiments followed GLP guidelines.

Blinding of drugs and/or spikes is indicated with the data.

BRD: WB/IL-6 March, 2006

5.6 Circumstances

Indicate the "lot-to-lot" consistency of the test substances, the time frame of the various studies, and the laboratory in which the study or studies were conducted. A coded designation for each laboratory is acceptable.
In each part of the study, all samples are derived from one (clinical) lot.

5.7 Other data available

Indicate the availability of any data not submitted for external audit, if requested.
All relevant data were submitted with the present BRD.

6 Test Method Accuracy

6.1 Accuracy

Describe the accuracy (e.g., concordance, sensitivity, specificity, positive and negative predictivity, false positive and negative rates) of the proposed test method compared with the reference test method. Explain how discordant results in the same or multiple laboratories from the proposed test were considered when calculating accuracy.

Test method accuracy was assessed in two large scale experiments performed with the drugs outlined in table 3.3.1 and table 3.3.2 in section 3 respectively. As described before one experiment was performed in an early stage of the study with 3 different drugs, tested 3 times and the other final experiment all drugs were tested once in the three participating laboratories. From the first experiment a preliminary estimate of sensitivity and specificity can be figure out, whereas the second is regarded as the established accuracy for the WB/IL-6 assay.

6.1.1 Preliminary estimate of the accuracy of the WB/IL-6 test. In an early stage of the study a different concept for interference testing was used. The developing laboratories determined for each drug (outlined in table 3.3.2, section 3.3) the smallest dilution within the MVD that showed no interference or an acceptable degree of interference with the spike recovery. In general the lowest dilution of the sample allowing for a 50-200% spike recovery was chosen. In addition, the positive control (PC) set at 0.5 EU/ml saline was used as the classification threshold. The laboratory procedure as described in the method protocol was maintained throughout the study. Although it was realized there were some drawbacks to the concept for interference testing and applying the PC as a threshold, this small-scale study allows for a preliminary estimate of the accuracy of the WB/IL-6 method.
It has to be noted that this part of the study was designed to provide an estimate of the intra- and interlaboratory reproducibility. Therefore it will also be discussed in detail in section 7 (Test Method Reliability).

According to the PM applied during an early phase of the study the outcome (positive/negative) is related to the positive control (PC=0.5 EU/ml). If absorbance of sample > absorbance of PC, then the sample is classified as positive. If absorbance of sample < PC, then the sample is classified as negative. While performing the experiments during this phase it was realized that there is a flaw between interference testing and the PM. After the interference testing, the lowest dilution of the sample allowing for a 50-200% spike recovery was chosen. It is obvious that even with a flawless repeatability of the assay; a spike recovery between 50%-100% would be classified as negative according to the preliminary PM. In addition, due to unforeseen problems with the preparations of the spike, the recovery of the spikes was far below 100%. (This is outside the scope of the study and will not be discussed). As a consequence of the employed preliminary setup of the study the sensitivity will be underestimated, and the specificity will be overestimated.

In short, three test substances were spiked with WHO-LPS at four levels (0, 0, 0.5 and 1.0 EU/ml respectively), which resulted in 12 samples with a balanced design with regard to positive and negative samples (i.e. samples expected to be pyrogenic and non-pyrogenic respectively. These 12 sample were three times tested in three laboratories. In total there were 108 classifications from 12 samples in 3 runs and in 3 laboratories (3x3x12=108). Results are described in detail in section 7. A 2x2 contingency table was constructed (table 6.1.1), from which the estimates of sensitivity and specificity can easily be derived.

Table 6.1.1: 2x2 contingency table. The prediction model applied to a preliminary study.

		True status of samples +	-	**Total**
PM	+	40	4	44
	-	14	50	64
	Total	54	54	108

The specifications of specificity and sensitivity described in section 5.3 were applied to these results and the specificity (Sp) of the WB/IL-6 assay is 93% (50/(4+50)*100%), 95% confidence interval [0.821; 0.979]. The sensitivity (Se) equals 74% (40/(40+14)*100%), 95% confidence interval [0.603; 0.850]. As outlined previously the specificity is overestimated and the sensitivity is underestimated as a result of the design of this part of the study.

6.1.2 Test method accuracy of the proposed WB/IL-6 method. To assess accuracy of the proposed method, 10 substances (listed in table 3.1.1, section 3) were spiked with five different concentrations of the WHO-LPS (one of which is negative). Thus, in total, 50 samples have been tested in each laboratory.
To permit the application of the chosen PM, each drug was diluted to its individual MVD, which has been calculated using the ELCto that drug (listed in section 3). Lesser dilutions were tested by the DL, and showed no interference. Therefore interference was not expected at the individual MVD. Each substance had their own specific set of endotoxin spike solution, ensuring that after spiking the undiluted drug, it contained 0.0; 0.25; 0.5; 0.5 and 1.0 EU/ml respectively when tested at its MVD. A detailed description of the sample preparation was supplied to the three independent laboratories (shown in table 5.1.1 for convenience). To put more weight to the result of this part of the validation, the spikes were blinded and coded by QA ECVAM. The raw data and the graphical presentation of these raw data are shown in the section 5 (table 5.4.2). Accuracy was assessed by applying the PM to the results (summarized in table 5.3.2) and evaluating the concordance in this section in a two by two contingency table (table 6.1.2). As described above 10 substances, spiked with 5 different WHO-LPS concentrations were tested in three laboratories and consequently a maximum of 150 data were available for analysis.

As intralaboratory reproducibility was successfully shown in previous experiments (analyzed in section 7), only one run performed in each laboratory was considered sufficient.

Table 6.1.2: 2x2 contingency table. Prediction model applied to the WB/IL-6 test result of 10 different substances assessed in three different laboratories.

		True status of samples +	-	Total
PM	+	79	2	81
	-	10	57	67
	Total	89	59	148

Of the 150 available data, two sets of data had to be removed from the analysis because the coding of the samples was mixed up by the testing laboratory. All quality criteria as defined in the method protocol were met. The specificity and sensitivity of the WB/IL-6 method could be estimated as described in section 5.3.

The specificity of the WB/IL-6 assay is 96.6% (57/(2+57)*100%), 95% confidence interval [0.883; 0.996]. The sensitivity equals 88.8% (79/(79+10) *100%), 95% confidence interval [0.803;0.945]. (See table 6.1.3). The specificity varied from 89% up to 100% within the three laboratories, and the sensitivity varied from 83% up to 100%.

Table 6.1.3: Specificity and sensitivity of the WB/IL-6 assay

	N total	N correctly identified	proportion	95% CI lower limit	95% CI upper limit
Specificity (Sp)	59	57	96.6%	88.3%	99.6%
Sensitivity (Se)	89	79	88.8%	80.3%	94.5%

6.2 Concordancy to *in vivo* reference method

Discuss results that are discordant with results from the in vivo reference method.
Not applicable.

6.3 Comparison with reference methods

Discuss the accuracy of the proposed test method compared to data or recognized toxicity from the species of interest (e.g., humans for human health-related toxicity testing), where such data or toxicity classification are available. This is essential when the method is measuring or predicting an endpoint for which there is no preexisting method. In instances where the proposed test method was discordant from the in vivo reference test method, describe the frequency of correct predictions of each test method compared to recognized toxicity information from the species of interest.
Not applicable.

6.4 Strength and limitations

State the strengths and limitations of the proposed test method, including those applicable to specific chemical classes or physical-chemical properties.
It appears the proposed test is applicable to most classes of medicinal products, at least those that are non- or low-toxic to cells *in vitro*. In addition, the test may be employed to

assess pyrogenicity of various medical devices, such as (biological) bovine collagen bone implants.

6.5 Data interpretation

Describe the salient issues of data interpretation, including why specific parameters were selected for inclusion.
No issues.

6.6 Comparison to other methods

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the results obtained with both test methods should be compared with each other and with the in vivo reference test method and/or toxicity information from the species of interest.
Not applicable.

7 Test Method Reliability (Repeatability/Reproducibility)

7.1 Selection of substances

Discuss the selection rationale for the substances used to evaluate the reliability (intralaboratory repeatability and intra- and interlaboratory reproducibility) of the proposed test method as well as the extent to which the chosen set of substances represents the range of possible test outcomes.

The rationale for the selection of the substances is described in section 3.3. In short: for the present studies endotoxin (WHO-LPS) was selected as the model pyrogen, since it is well defined biological standard and readily available. Selected test substances were medicinal products available on the market. These batches are released by the manufacturers and comply with the Marketing Authorisation file and European Pharmacopoea. Therefore these batches are considered to contain no *detectable* pyrogens. To test the method reliability the medical products were spiked with endotoxin.

7.2 Results

Provide analyses and conclusions reached regarding the repeatability and reproducibility of the proposed test method. Acceptable methods of analyses might include those described in ASTM E691-92 (13) or by coefficient of variation analysis.

In an early phase of the study, the intralaboratory repeatability and reproducibility of the test method was assessed in a series of experiments conducted in the DL. Series of blanks (saline, 0 EU/ml) and spikes of WHO-LPS in saline were tested. These experiments (1A, 1B, 2B and 2C) are summarized in table 7.2.1.

Table 7.2.1: Summary of experiments with WHO-LPS in saline.

Experiment	Spikes (EU/ml) in saline	n (per spike)	Repetitions of experiment	N
1A	0; 0.5	20	1	40
1B	0; 0.063; 0.125; 0.25; 0;5	10	1	50
2B	0; 0.25; 0.5	8	3	72
2C	0; 0.5	5	8	80

The data were used to answer questions regarding the nature of the distribution, the variance and its behavior over the range of response in replicated measurements under identical conditions. In addition, reliability of the test method was assessed by the maintenance of minimum criteria, e.g. a detection limit below an a-priori chosen positive control or a dose-dependent standard curve (table 7.2.1, experiment 1B). With the data of this experiment an assessment of the limit of detection of the test can be done by calculating the smallest spike, which can be discriminated from the blank.

The second group of experiments was meant to analyze the variation in detail. For this purpose the major sources of variation were assessed separately, i.e. behavior of a donor in time (experiment 2A), operator (exp. 2B) different donors (exp.2C). A total of 242 data were collected and analyzed.

First the shape of the distribution at a spike was assessed (not shown). Most of the data showed normal-distribution.
Based on the experience that there is a monotone increasing relationship between the mean-responses and the variation (empirical variance or standard deviation), the analysis focuses on the coefficient of variation (CV). The CV should be distributed symmetric around a constant factor, if the mean-variance relationship is linear. A plot of all CVs against their corresponding means is shown in figure 5.2.1. From the figure it is clear that at this stage of the study, the CV for some sets of replicates of the blanks is exceptionally high with CV 0.5 and 0.9. (From subsequent studies it appeared that this high variation of the blanks was just an incident). For the spikes with WHO-LPS, the variation for the sets of replicates is low. As only WHO-LPS was examined up to this point, it was envisaged that the CV would increase with other substances being tested. For CV criteria applied as a validity criteria of the WB/IL-6 assays, the CV was arbitrarily set at CV<0.4.

The outliers were identified on the assumption of normally distributed data as well as a log-normal distribution. At this point the Grubbs-test was chosen and the kind of outlier (lower or upper) and the significance level α (5% and 1% significance level) were recorded. Altogether there were 11and 8 outliers identified for the assumption of normality and log-normality respectively. Overall the amount of outliers is about 3.5-4%. The outliers were located all over the ELISA-plates and there was no obvious scheme. In addition, the raw data (plate-readouts) showed no obvious edge-effects or trends.

The results of test 1A (figure 5.2.2) show a low variation and the spike of 0.5 EU/ml was clearly detected. Test 1B showed a higher variation and the 0.25 EU/ml spike hardly discriminated from the blank (figure 5.2.2, one outlier for the blank, one outlier for the 0.063 EU/ml-spike). However the highest spike (0.5 EU/ml) can be detected easily.

Test 2A was included to assess the behavior of a donor in time. The blood of one donor was employed twice on different dates. Data are presented in figure 5.2.3. In both experiments the response of the donor are similar. In comparison with the (higher) variation between different donors (test 2C), the variation of a suitable donor is low and is considered to be no critical issue in the WB/IL-6 assay.

Experiment 2B (figure 5.2.4) was conducted by three operators in parallel with blood from one donor. Every operator tested eight replicates of three spikes of 0, 0.25 and 0.5 EU/ml-LPS. Obviously the operator has an impact on the results and the variability of the replicates seems to depend on (the experience of) the operator. Still, the data of the 0.5 EU/ml spikes can be discriminated from their corresponding blanks (after removal of the outliers).

The final experiment was designed to show the robustness of the assay with respect to different donors. Therefore 8 donors were involved and for each donor five replicates of each of the spikes (0; 0.5EU/ml) were generated. Data are presented in figure 5.2.5. Some variation in sensitivity for LPS between the donors is obvious. But every donor reacts to

the 0.5 EU/ml-spike. This experiment reveals that there is a certain effect of the covariate "donor" which is however not crucial, at least not with regard to a qualitative PM.

In conclusion: The most critical issue identified is the variation within the sets of blanks, but this is probably caused by the handling of the assay. The WB/IL-6 assay is robust against all examined variables. Although the experiments revealed an effect for the covariates "blood donor", "operator" and "day", the sensitivity of the assay is about 0.25 EU/ml and at least 0.5 EU/ml for all experiments. Therefore the intralaboratory repeatability is considered satisfactory. The 3-4% percentage outliers, as determined by the Grubbs test is considered acceptable. The validity criteria of the WB/IL-6-assay as recorded in the method protocol, are based on these experiments, i.e. CV< 0.4, lower limit of detection 0.5 EU/ml.

Intra- and interlaboratory reproducibility.
After transfer of the WB/IL-6 assays to two other laboratories, a dose response experiments was performed by all three laboratories. For this study 7 concentrations were tested in a dose response curve (0, 0.125, 0.25, 0.5, 1, 2, 4 EU/ml, at least 8 replicates). A participating laboratory qualified for taking part in the next part of the study by producing a dose response curve, with a limit of detection of at least 0.5 EU/ml and a CV < 0.4 (data not shown).

The intra- and interlaboratory reproducibility was assessed by testing 3 different medicinal substances, Gelafundin, Jonosteril and Haemate (described in table 3.3.2, section 3.3.). Test substances and their spikes were appropriately blinded. Test substances were tested, at a predefined dilution above the MVD, independently in 3 different laboratories, 3 times each. The three test substances were spiked with WHO-LPS at four levels (0, 0, 0.5 and 1.0 EU/ml respectively), which resulted in 12 samples with a balanced design with regard to positive and negative samples (i.e. samples expected to be pyrogenic and non-pyrogenic respectively. In addition a negative control (saline) and positive control (0.5 EU/ml in saline) were included to establish assay validity. To avoid interference, the DL performed interference testing in terms of the BET, i.e. 50-200% spike recovery, and decided on the dilution of the test substances. Dilutions chosen for Gelafundine, Haemate, Jonosteril were 1:2, 1:4 and 1:1 respectively. The results are graphically presented using the absorbance values of the three runs (shown in section 5, fig. 5.2.5).

From the experiment with LPS-WHO only it was concluded that CV for the WB/IL-6 assay is < 0.4, which is acceptable. It was envisaged that the CV was likely to be higher when testing different substances (different matrices) and was assessed for the current set of data. A plot of all CVs for all sets of 4 replicates of a drug with a spike is shown in figure 5.2.7. From the figure it is clear that the CV for a set of 4 replicates of one spike concentration is usually below 0.45, which is considered acceptable for a biological assay. Only one set of data showed an exceptional high (CV>1.1) which is probably due to a pipetting error. For the remainder of the studies the CV criteria applied as validity criteria of the WB/IL-6 assays was arbitrarily set at CV<0.45.

BRD: WB/IL-6 March, 2006

The analysis of the intralaboratory reproducibility was assessed from the three identical and independent runs conducted in each laboratory. The comparison of the three runs was carried out blindly such that the laboratory did not know the true classification of the sample (either pyrogenic or non-pyrogenic). By this procedure only the randomness and reproducibility of the methods was assessed and not systematic errors, which may have arisen from other sources, e.g. logistics or preparation of the samples. This means that if a sample was misclassified in all three runs the result is 100% intralaboratoryoratory reproducible (regardless of the misclassification of the sample).

According to the preliminary PM applied during this phase of the study the outcome (positive/negative) was related to the positive control (PC=0.5 EU/ml). If absorbance of sample > absorbance of PC, then sample is classified as being positive. If absorbance of sample < PC, sample is classified as negative (positive/pyrogenic = 1, negative/non-pyrogenic = 0).
During this phase it was realized that there is a flaw between interference testing and the PM. After the interference testing, the lowest dilution of the sample allowing 50-200% spike recovery was chosen. It is obvious that even with a flawless repeatability of the assay, a spike recovery between 50%-100% would be classified as negative according to the preliminary PM.

From the three independent runs summarized in table 5.4.1, the intralaboratory reproducibility can be calculated for the separate laboratories (table 7.2.2). For these calculations there is no need for information of the true status of the sample. A minimum criterion for the establishment of an assay is that experiments carried out with the same samples should result in a high concordance of classifications.
Each of the assays performed by the laboratories fulfilled the sensitivity criterion, i.e. the assays showed a significant difference between C- and C+. All results could be included in the analysis. From table 7.2.2 it can be read that the between runs reproducibility ranges from 75 to 100%. The mean intralaboratory reproducibility is very good (83 – 100%) for all three participating laboratories.

Table 7.2.2 : Intralaboratory reproducibility, assessed by correlation between different runs. Result of testing 3 substances 3 times by 3 laboratories.

	DL (NIBSC)	**NL1 (Basel)**	**NL2 (Innsbruck)**
Run 1 - Run 2	75% (9/12)	92% (11/12)	100% (12/12)
Run 1 - Run 3	100% (12/12)	92% (11/12)	100% (12/12)
Run 2 - Run 3	75% (9/12)	100% (12/12)	100% (12/12)
Mean	83%	94%	100%
Proportion showing the same result in 3 runs	75%	92%	100%

BRD: WB/IL-6 March, 2006

The interlaboratory reproducibility of the WB/IL-6 method was assessed in a similar manner to the intralaboratory reproducibility. A summarizing method to combine the three runs per laboratory is considered not appropriate, because it would mask misclassification. Therefore each run of one laboratory was compared with all runs of another laboratory. This results optimally in 108 comparisons between the data sets of two laboratories. The measure of similarity is then the proportion of equally classified samples. These proportions are summarized in table 7.2.3, show that there is a good interlaboratory reproducibility varying from 72 - 97% (overall mean: 81%).

Table 7.2.3: Interlaboratory reproducibility. Assessed by interlaboratory correlations. Result of testing 3 substances 3 times by 3 laboratories.

Laboratories	Interlaboratory Reproducibility	Number of equal predictions
DL – NL1	72%	78 / 108
DL – NL2	75%	81 / 108
NL1 – NL 2	97%	105 / 108
Mean	81%	

DL = NIBSC; NL1 = Basel; NL2 = Innsbruck

Also from the result of the large-scale study (testing 10 substances spiked with 5 separate spikes), the interlaboratory reproducibility can be estimated (table 7.2.4). The reproducibility varied from 85% to 88% between two laboratories. All three laboratories found the same result for 38 samples out of 48 (equals 79%).

Table 7.2.4: Interlaboratory reproducibility. Assessed by testing of 10 substances, spiked 5 times. One run of 50 samples by three different laboratories.

Laboratories	Interlaboratory Reproducibility	Number of equal predictions
DL - NL1	85%	41 / 48
DL - NL2	85%	41 / 48
NL1 – NL2	88%	44 / 50
Mean	86%	
same result in all laboratories	79%	38 / 48

DL =PEI; NL1 = Basel; NL2 = Innsbruck

Conclusion: It is shown that the mean intralaboratory reproducibility, assessed by the proportion of equally classified samples between different runs varies from 83% to 100% between the three participating laboratories. The interlaboratory reproducibility between two laboratories varied from 72% to 97% in one large scale blinded experiment and from 85% to 88% in the other large scale blinded experiment. All three participating laboratories predicted the same in 79% of the measurements. It has to be noted that part of the samples was 0.5 EU/ml and close to the arbitrary point of the WB/IL-6 assay.

BRD: WB/IL-6 March, 2006

7.3 Historical data

Summarize historical positive and negative control data, including number of experiments, measures of central tendency, and variability.
Not applicable.

7.4 Comparison to other methods

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the reliability of the two test methods should be compared and any differences discussed.
Not applicable.

8 Test Method Data Quality

8.1 Conformity

State the extent of adherence to national and international GLP guidelines (7-12) for all submitted data, including that for the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method. Information regarding the use of coded chemicals and coded testing should be included.
The studies were done in accordance to the guidelines for GLP. Written protocols and approved standard operating procedures were followed during the entire course of the study. Deviations were recorded and, where appropriate, approved in amendments. All data are stored and archived. As mentioned, samples were appropriately blinded.

8.2 Audits

Summarize the results of any data quality audits, if conducted.
No audits were conducted.

8.3 Deviations

Discuss the impact of deviations from GLP guidelines or any noncompliance detected in the data quality audits.
Not applicable.

8.4 Raw data

Address the availability of laboratory notebooks or other records for an independent audit. Unpublished data should be supported by laboratory notebooks.
All records are stored and archived by the contributing laboratories and available for inspection.

9 Other Scientific Reports and Reviews

9.1 Summary

Summarize all available and relevant data from other published or unpublished studies conducted using the proposed test method.

Relevant data obtained with the proposed method are described in a number of published studies which are given in Appendix B. The most important results will be summarized below.

An *in vitro* monocyte activation test that detected pro-inflammatory and pyrogenic contaminants, was first applied some 15 years ago (Poole et al., 1988). A number of variants of the original test system have since been described, although the underlying principle of each variant remains the same. The test preparation is cultured with monocytes, either as peripheral blood mononuclear cells, PBMC, diluted whole blood or cells of a monocytoid cell line such as MM6. Contaminants in the test article activate CD14/TLR receptors which stimulates the release of an endogenous pyrogenic cytokine from the monocytes (Poole and Gaines Das, 2001).
Early studies mainly report on optimization of the test method, e.g. improving the lower limit of detection, incubation times and cytokine readout, using model pyrogens such as LPS or endotoxin. Only limited information is available on the actual testing of medicinal products.

Most interestingly, Taktak et al (1991) described several batches of a medicinal product (serum albumin) that caused adverse (pyrogenic) reactions in recipients. These lots were not detected by either BET or rabbit test but only by the *in vitro* monocytoid cell test.
In a study using whole blood and monocytoid cell lines as the sources of monocytoid cells (Nakagawa et al., 2002) it was reported that the structurally diverse pyrogens endotoxin, peptidoglycan, Staphylococcus aureus Cowan 1 and poly(I.C) all stimulated the release of cytokines.

The cytokine readout included tumor necrosis factor (TNF), interleukin-1 (IL-1), and IL-6 (reviewed by Poole and Gaines Das, 2001 and Poole et al., 2003). Other cytokines, e.g. IL-8, are also produced in large quantities in response to pyrogenic contaminants but their roles in fever are less well studied. The preferred readout is usually IL-6 because IL-6, unlike IL-1 and TNF, is secreted entirely into the cell-conditioned medium in large quantities, thereby permitting its complete estimation (Poole et al, 1988; Poole et al., 1989 and Taktak et al., 1991). No significant differences were observed in the kinetics or production levels of IL-6 in whole blood and PBMC (de Groote et al.,1992)
Nakagawa et al. (2002) compared TNF alfa, IL-1b and IL-6 as readout, with diluted whole blood and a monocytoid cell line (MM6, clone CA8). The structurally diverse pyrogens endotoxin, peptidoglycan, S. aureus Cowan 1 and poly (IC) all stimulated the release of more IL-6 than either TNF or IL-1. More importantly, IL-6 was induced by lower concentrations of each pyrogen.

It has also been shown that certain pro-inflammatory bacterial components stimulate the production of IL-6 but not TNF and IL-1 (Reddi et al., 1996), and IL-6 induction via Toll-like (pyrogen) receptors rapidly follows the recognition of microbial products (Pasare and Medzhitov, 2003).

It has been recognized before that for the routine applications of cytokine release tests the simplicity of the whole blood method is more suitable than a test with isolated monocytes (Schins 1996). Therefore many researchers have focused on the whole blood method.

Good correlation were found between the WB/IL6 assay and the rabbit pyrogen assay for 22 freshly prepared production batches of human serum albumin, fibronectin and stabilized human serum solutions. None of the products had an effect on the sensitivity of the WB/IL6 assay whereas the BET gave anomalous results for 1 out the 22 production batches tested. (Pool, 1998)

Even a strategy to differentiate between endotoxin and non-endotoxin pyrogens, using Polymyxin B has been suggested (Pool, 1999). Relatively high concentrations of Polymyxin B inhibits endotoxin-induced IL-6 secretion by whole blood cells. Polymyxin B could partially inhibit IL-6 induction by 2 batches of HSA that were highly pyrogenic using the rabbit and the whole blood assay, suggesting that non-endotoxin pyrogens were present. This was also supported by BET result, showing only a weak positive or inconclusive result for these batches. However, this challenging idea to differentiate between endotoxin and non-endotoxin needs further research.

It is stressed throughout these studies using whole blood that only healthy donors not taking any medication must be used for testing pharmaceuticals. Various drugs such as cortisone suppress interleukin release and would therefore exclude a blood donor from the experiment. An infection can stimulate release, as reflected in an increased baseline value or an abnormally high interleukin response. Therefore, the WB/IL-6 test may only be used if samples have first been shown not to cause interference. The blood group of the human donors does not influence the results of the assay.

The pyrogenicity of a complex multivalent vaccine, Infanrix, containing protein and polysaccharide components from both gram-positive and gram-negative bacteria, was studied using the WB/IL6 test. The study revealed a large variability in IL6 production by different donors. Although all blood samples responded to endotoxin, only some donors significantly responded to Infanrix. (The blood donors histories of vaccinations and infections were not recorded). Infanrix was negative in the BET, but interfered with the spike recovery of endotoxin. The significance of this finding with such a complex mixture as a multivalent vaccine remains to be elucidated.

9.2 Discussion

Comment on and compare the conclusions published in independent peer-reviewed reports or other independent scientific reviews of the proposed test method. The conclusions of such scientific reports and reviews should be compared to the conclusions reached in this submission. Any ongoing evaluations of the proposed test method should be described.

The validation described in this BRD is the first time such an extensive study for specificity and accuracy using actual medicinal products spiked with endotoxin is carried out. Moreover it is the first time that similar samples were tested in parallel by (at least) three laboratories. Although the laboratories had not the same level of hands on with this particular assay, the accuracy is comparable between the three laboratories. There are no reports in independent peer-reviewed journals available to compare the accuracy in multiple laboratories, except for the manuscript of Hoffman et al (2005b). It is shown that validated assays employing, either MM6 cells, whole blood or PBMC have comparable accuracies.

Quite recently a comparative evaluation for two different *in vitro* tests for pyrogens, using PBMC and diluted whole blood respectively, was published (Andrade et al. 2003). Both tests, with a IL6 readout, were applied to different classes of parenteral medicinal products. Many of these products did not have a specified ELC that was established as the MVD to comply with the test. Preparatory tests were conducted to ensure that the drugs being tested did not interfere in the tests. Both *in vitro* tests showed a good overall agreement, both with each other and with the BET and the rabbit pyrogen test for the detection of endotoxin. The batch of medicinal product failing the rabbit test, was also positive in BET, whole blood and PBMC test. In addition, the whole blood test was shown to be sensitive to the fungus *C. albicans* and the gram-positive bacteria *S. aureus*.

9.3 Results of similar validated method

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the results of studies conducted with the validated test method subsequent to the ICCVAM evaluation should be included and any impact on the reliability and accuracy of the proposed test method should be discussed.

As mentioned, *in vitro* monocytoid activation test methods for the detection of pyrogenic contaminants are being developed over the course of the past two decades. A number of variants have been described, although the underlying principle of each variant remains the same. The test preparation is cultured with monocytoid cells, either as peripheral blood mononuclear cells, PBMC, (diluted) whole blood or cells of a monocytoid cell line such as MM6. Accuracy and specificity of these test methods are comparable, but in general methods using whole blood, PBMC and the MM6 cell line appear to perform best (Hoffmann et al, 2005b).

Table 9.3.1 summarises the performance of *in vitro* methods presented in the five BRDs and Table 9.3.2 compares the *in vivo* and *in vitro* pyrogen tests regarding their strengths, weaknesses, costs, time, limitations.

However, most studies (as this one) are done with model pyrogens and as yet little experience is available in the field, e.g. as part of the final batch release test-package. Experience and thus confidence in these methods will grow once regulatory authorities approve these methods and more manufacturers start to employ them. Then, on a case by case situation, it should be determined which method is best suited for the actual situation and demonstrates to pick out the appropriate, i.e. pyrogenic batches of the medicinal product.

Table 9.3.1: Summary of the performance of in vitro pyrogen tests based on monocytoid cells (see Tables 7.2.2; 7.2.4; 6.1.3)

Test	System	Read-out	Intralaboratory reproducibility (%)	Interlaboratory reproducibility (%)	Sensitivity (%)	Specificity (%)
WB/IL-6	whole blood	IL-6	DL: 83.3 NL1: 94.4 NL2: 100	DL-NL1: 85.4 DL-NL2: 85.4 NL1-NL2: 92.0	88.9	96.6
WB/IL-1	whole blood	IL-1β	DL: 88. 9 NL1: 95.8 NL2: 94.4	DL-NL1: 72.9 DL-NL2: 81.6 NL1-NL2: 70.2	72.7	93.2
96-wells WB/IL-1 [1]	whole blood	IL-1β	-	DL-NL1: 88.1 DL-NL2: 89.7 NL1-NL2: 91.5	98.8	83.6
CRYO WB/Il-1	cryo whole blood	IL-1β	-	DL-NL1: 91.7 DL-NL2: 91.7 NL1-NL2: 91.7	97.4	81.4
KN CRYO WB/Il-1 [2]	cryo whole blood	IL-1β	-	DL-NL1: 83.3 DL-NL2: 100 NL1-NL2: 83.3	88.9	94.4
PBMC/IL6	PBMC	IL-6	DL: 94.4 NL1: 100 NL2: 94.4	DL-NL1: 84.0 DL-NL2: 86.0 NL1-NL2: 90.0	92.2	95.0
PBMC-CRYO/IL-6 [3]	PBMC	IL-6	-	DL-NL1: 96 DL-NL2: 76 NL1-NL2: 80	93.3	76.7
MM6/IL-6	MM6	IL-6	DL: 100 NL1: 94.4 NL2: 94.4	DL-NL1: 90.0 DL-NL2: 89.6 NL1-NL2: 83.3	95.5	89.8

DL = developing laboratory; NL1, NL2 = naive laboratory 1 and 2
1 = data provided in Section 13 of WB/IL-1 BRD
2 = data provided in Section 13 of CRYO WB/IL-1 BRD
3 = data provided in Section 13 of PBMC/IL-6 BRD
Table amended from Hoffmann et al 2005b; results with THP cells not included

BRD: WB/IL-6 March, 2006

Table 9.3.2: Comparison of the in vivo and in vitro pyrogen tests regarding their strengths, weaknesses, costs, time, limitations

	Rabbit pyrogen test	**BET / LAL**	**In vitro pyrogen test**
Test materials	Liquids	Clear liquids	Liquids, potentially cell preparations, solid materials
Pyrogens covered	All (possible species differences to humans for non-endotoxin pyrogens)	Endotoxin from Gram-negative bacteria	(probably) all
Limit of detection (LPS)	0,5 EU	0,1 EU (some variants down to 0,01 EU)	0,5 EU (validated PM), some variants down to 0,001 EU
Ethical concerns	Animal experiment	About 10% lethality to bled horseshoe crabs	Some assays: blood donation
Costs*	High (200-600$/sample)	Low (50-150$/sample)	Medium (100-350$/sample)
Time required	27 h	45 min	24-30h**
Materials not testable	Short-lived radiochemicals, anesthetics, sedatives, analgetics, chemotherapeutics, immunomodulators, cytokines, corticosteroids	Most biologicals, glucan-containing preparations (herbal medicinal products, cellulose-filtered products), lipids, microsomes, cellular therapeutics	Not known (some of the materials not testable in rabbits require adaptations)
Others	No positive or negative control included, strain differences, stress affects body temperature	Potency of LPS from different bacterial species in mammals not reflected, false-positive for glucans	Possible donor differences, need to exclude hepatitis/HIV and acute infections / allergies of donors, dedifferentiation of cell lines

* = We consulted the laboratories participating in the validation study and a consultant regarding the costs of the tests. The figures we received vary significantly depending on the facility (e.g. industry, contract laboratory, control authority), frequency of testing, specific test requirements, country, etc.

** = interference testing might increase duration by 24 hours

10 Animal Welfare Considerations (Refinement, Reduction, and Replacement)

10.1 Diminish animal use

Describe how the proposed test method will refine (reduce or eliminate pain or distress), reduce, or replace animal use compared to the reference test method.

Depending on the medicinal product, one of two animal-based pyrogen tests is currently prescribed by the respective Pharmacopoeias, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET). The rabbit pyrogen test detects various pyrogens but alone the fact that large numbers of animals are required to identify a few batches of pyrogen-containing samples argues against its use. In the past two decades, the declared intention to refine, reduce and replace animal testing, has lowered rabbit pyrogen testing by 80% by allowing to use the BET as an alternative pyrogen test for certain medicinal products. Bacterial endotoxin is the pyrogen of major concern to the pharmaceutical industry due to its ubiquitous sources, its stability and its high pyrogenicity. With the BET, endotoxin is detected by its capacity to coagulate the amoebocyte lysate from the haemolymph of the American horseshoe crab, *Limulus polyphemus*, a principle recognised some 40 years ago (Levin and Bang, 1964). In the US, *Limulus* crabs are generally released into nature after drawing about 20% of their blood and therefore most of these animals survive. However, the procedure still causes mortality of about 30,000 horseshoe crabs per year, which adds to more efficient threats of the horseshoe crab population like its use as bait for fisheries, habitat loss and pollution.

The proposed test method is an alternative for the rabbit test and the BET. By replacing the rabbit test or the BET, the lives of rabbits and horseshoe crabs are spared.

10.2 Continuation of animal use

If the proposed test method requires the use of animals, the following items should be addressed:

10.2.1 *Describe the rationale for the need to use animals and describe why the information provided by the proposed test method requires the use of animals (i.e., cannot be obtained using non-animal methods).*

Not applicable.

10.2.2 *Include a description of the sources used to determine the availability of alternative test methods that might further refine, reduce, or replace animal use for this testing. This should, at a minimum, include the databases searched, the search strategy used, the search date(s), a discussion of the results of the search, and the rationale for not incorporating available alternative methods.*

Not applicable.

10.2.3 *Describe the basis for determining that the number of animals used is appropriate.*

Not applicable.

BRD: WB/IL-6 March, 2006

10.2.4 *If the proposed test method involves potential animal pain and distress, discuss the methods and approaches that have been incorporated to minimize and, whenever possible, eliminate the occurrence of such pain and distress.*
Not applicable.

11 Practical Considerations

11.1 Transferability

Discuss the following aspects of proposed test method transferability. Include an explanation of how this compares to the transferability of the in vivo reference test method and, if applicable, to a comparable validated test method with established performance standards.

In general, the proposed test method is not unlike other bioassays and immunoassays that are performed routinely in many laboratories.

11.1.1 *Discuss the facilities and major fixed equipment needed to conduct a study using the proposed test method*.

No extraordinary facilities are required. General laboratory equipment and analytical instruments for performing immunoassays, e.g. microtiter plate reader and –washer, are sufficient to perform the proposed test method.

11.1.2 *Discuss the general availability of other necessary equipment and supplies*.

All supplies and reagents are readily available on the market. In contrast, availability of sufficient rabbits of adequate weight and in good health for the *in vivo* reference test is sometimes reported a limitation.

It is stressed throughout these studies using whole blood that only healthy donors not taking any medication must be used for testing pharmaceuticals. Various drugs such as cortisone suppress interleukin release and would therefore exclude a blood donor from the experiment. An infection can stimulate release, as reflected in an increased baseline value or an abnormally high interleukin response.

11.2 Training

Discuss the following aspects of proposed test method training. Include an explanation of how this compares to the level of training required to conduct the in vivo reference test method and, if applicable, a comparable validated test method with established performance standards.

11.2.1 *Discuss the required level of training and expertise needed for personnel to conduct the proposed test method.*

The proposed test method requires personnel trained for general laboratory activities in cell biology and immunochemistry or biochemistry. Techniques they should master are not unlike cell culture (aseptic operations) and immunological techniques (especially ELISA). Such expertise is available in most if not all QC-laboratories.

11.2.2 *Indicate any training requirements needed for personnel to demonstrate proficiency and describe any laboratory proficiency criteria that should be met.*

Personnel should demonstrate that they master the execution of the test. The candidate should demonstrate to meet all the appropriate assay acceptance criteria and yield accurate results (outcome) using selected test items.

11.3 Cost Considerations

Discuss the cost involved in conducting a study with the proposed test method. Discuss how this compares to the cost of the in vivo reference test method and, if applicable, with that of a comparable validated test method with established performance standards.
Two factors contribute to the cost of the proposed test method: cost of the reagents and especially personnel.
Since the proposed test method is relatively more labor-intensive, it is estimated that the cost of the proposed test method is more then the BET or the *in vivo* reference test using rabbits. Obviously, a higher throughput of tests (runs/year) such as in a QC-laboratory of a multi-product facility or in a Contract Research Organization will significantly reduce the costs per assay.
However, especially with pharmaceuticals of biological origin, the proposed test method may be cost-effective, since these products all to often are incompatible with the BET and by their nature preclude the re-use of the rabbits.

11.4 Time Considerations

Indicate the amount of time needed to conduct a study using the proposed test method and discuss how this compares with the in vivo reference test method and, if applicable, with that of a comparable validated test method with established performance standards.
Essentially the test stretches two working days. On day one the testing materials are prepared and incubated overnight with the monocytoid cells. On the second day the amount of excreted cytokines is determined by immunoassay. The total time from start to result is approximately 24 hours.
It is thus concluded that the proposed test method will take more time when compared to the reference tests, either the rabbit test or the BET. It should be noted that rabbits are tested prior to their first use by a sham test.

12 References

List all publications referenced in the submission.

Andrade SS, Silveira RL, Schmidt CA, Junior LB, Dalmora SL. (2003) Comparative evaluation of the human whole blood and human peripheral blood monocyte tests for pyrogens. Int J Pharm. Oct 20;265(1-2):115-24.

Barnett, V., Lewis, T. (1984). Outliers in statistical data. 2nd edition, Chichester, US: John Wiley & Sons, pp 171-172.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Carlin G, Viitanen E. In vitro pyrogenicity of a multivalent vaccine: Infanrix. PHARMEUROPA, 2003, 15(3), 418-423.

Clopper, C. J., Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26: 404-413

Co Tui & Schrift, M.H. (1942). A tentative test for pyrogen in infusion fluids. Proc. Soc. Exp. Biol. Med. 49: 320-330.

Cooper JF, Levin J, Wagner Jr. HN (1971). Quantitative comparison of *in vitro* and *in vivo* methods for the detection of endotoxin. J Lab Clin Med 78: 138-145.

De Groote, D., Zangerle, P.F., Gevaert, Y., Fassotte, M.F., Beguin, Y., Noizat-Pirenne, F., Pirenne, J., Gathy, R., Lopez, M., Dehart I., (1992). Direct stimulation of cytokines (IL-1 beta, TNF-alpha, IL-6, IL-2, IFN-gamma and GM-CSF) in whole blood. I. Comparison with isolated PBMC stimulation. Cytokine 4, 239.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005a). Optimisation of pyrogen testing in parenterals according to different

pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Pertbauer A, Schindler S, et al. (2005b) Validation of novel pyrogen tests based on human monocytoid cell lines. J Immunol Methods, 298(1-2):161-173.

Hothorn, L. A. (1995). Biostatistical analysis of the 'control vs. k treatments' design including a positive control group. In: Testing principls in clinical and preclinical trials, pp 19-26, Stuttgart, Germany: Gustav Fischer Verlag.

Levin J, Bang FB (1964). A description of cellular coagulation in the Limulus. Bull Johns Hopkins Hosp 155: 337-345.

Nakagawa, Y., Maeda, H. and Murai, T. (2002). Evaluation of the *in vitro* pyrogen test system based on proinflammatory cytokine release from human monocytes: comparison with a human whole blood culture test system and with the rabbit pyrogen test. Clin. Diagn. Lab. Immunol. 9: 588-592

Pasare, C. and Medzhitov, R. (2003). Toll pathway-dependent blockade of CD4+CD25+ T cell-mediated suppression by dendritic cells. Science 299: 1033-1036.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1998) The detection of pyrogens in blood products using an ex vivo whole blood culture assay. J Immunoassay. May-Aug;19(2-3):95-111.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1999) Differentiation between endotoxin and non-endotoxin pyrogens in human albumin solutions using an ex vivo whole blood culture assay. J Immunoassay. Feb-May; 20(1-2):79-89.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, november 1989: 17-18.

Poole, S. and Gaines Das, R.E. (2001). Towards a 'human pyrogen test'. J. Parenter. Sci. 6: 63-66.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' *in vitro* test for pyrogens. J Immunol Methods 274: 209-220.

Reddi, K., Nair, S.P. et al (1996). Surface-associated material from the bacterium Actinobacillus actinomycetemcomitans contains a peptide which, in contrast to lipopolysaccharide, directly stimulates fibroblast interleukin-6 gene transcription. Eur. J. Biochem. 236: 871-876.

Snedecor, G. W., Cochran, W. G. (1989). Statistical methods, 8th edition, Ames, US: The Iowa State University Press.

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

13 Supporting Materials (Appendices)

13.1 Standard operating procedure (SOP) of the proposed method

Provide the complete, detailed protocol for the proposed test method.

Appendix A includes the complete and detailed protocol of the WB/IL-6 method (*Detailed protocol WB/Il-6 In vitro test for pyrogen/endotoxin using human whole blood 22 07 02; electronic file name: SOP WB IL6*) as used throughout the studies described in section 5 of the current BRD and the trial plan of the validation study.

13.2 Standard operating Procedure (SOP) of the reference method

Provide the detailed protocol(s) used to generate reference data for this submission and any protocols used to generate validation data that differ from the proposed protocol.
Not applicable.

13.3 Publications

Provide copies of all relevant publications, including those containing data from the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method with established performance standards.

Not all of the publications cited in the BRD (see section 12) are included as hardcopies in Appendix B, e.g. publications on statistical methods are not given.

Remark: The same set of hardcopies was included into Appendix B of all of the 5 submitted BRDs. Therefore some of the publications in Appendix B might not be referenced in the current BRD nor included in the list of the publications in section 12. Three general publications were added, which are not cited in any BRD but might be useful as background information to the validation study: the ECVAM publications on validation (Balls et al, 1995; Curren et al, 1995; Hartung et al, 2004). Several publications were included, which either give more background information on the human fever reaction or report on specific studies using *in vitro* pyrogen tests.

List of hard copies

Andrade SS, Silveira RL, Schmidt CA, Junior LB, Dalmora SL. (2003) Comparative evaluation of the human whole blood and human peripheral blood monocyte tests for pyrogens. Int J Pharm. Oct 20;265(1-2):115-24.

Balls et al. (1995) Practical aspects of the validation of toxicity test procedures. ECVAM Workshop Report 5. ATLA 23, 129-147.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Carlin G, Viitanen E. (2003) In vitro pyrogenicity of a multivalent vaccine: Infanrix. PHARMEUROPA, 15(3), 418-423.

Curren et al (1995) The Role of prevalidation in the development, validation and acceptance of Alternative Methods. ECVAM Prevalidation Task Force Report 1. ATLA 23, 211-217.

De Groote, D., Zangerle, P.F., Gevaert, Y., Fassotte, M.F., Beguin, Y., Noizat-Pirenne, F., Pirenne, J., Gathy, R., Lopez, M., Dehart I., (1992). Direct stimulation of cytokines (IL-1 beta, TNF-alpha, IL-6, IL-2, IFN-gamma and GM-CSF) in whole blood. I. Comparison with isolated PBMC stimulation. Cytokine 4, 239.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Fennrich S, Wendel A and Hartung T. (1999). New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 16:146-149

Fennrich S, Fischer M, Hartung T, Lexa P, Montag-Lessing T, Sonntag H-G, Weigandt M und Wendel A. (1999). Detection of endotoxins and other pyrogens using human whole blood. Dev. Biol. Standards 101:131-139

Gaines Das et al (2004). Monocyte activation test for pro-inflammatory and pyrogenic contaminants of parenteral drugs: test design and data analysis. J. of Immunological Methods 288, 165-177.

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hartung et al (2004) A modular approach to the ECVAM principles on test validity. ATLA 32, 467-472.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005). Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Peterbauer A, Schindler S, et al (2005). International validation of novel pyrogen tests based on human monocytoid cells. J Immunol Methods 298: 161-173.

Jahnke M, Weigand, Sonntag H-G. (2000). Comparative testing for pyrogens in parenteral drugs using the human whole blood pyrogen test, the rabbit in vivo pyrogen test and the LAL test. European Journal of Parenteral Science 5(2):39-44

Nakagawa Y, Maeda H, Murai T. (2002) Evaluation of the in vitro pyrogen test system based on proinflammatory cytokine release from human monocytes: comparison with a human whole blood culture test system and with the rabbit pyrogen test. Clin Diagn Lab Immunol. May;9(3):588-97.

Pasare, C. and Medzhitov, R. (2003). Toll pathway-dependent blockade of CD4+CD25+ T cell-mediated suppression by dendritic cells. Science 299: 1033-1036.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1998) The detection of pyrogens in blood products using an ex vivo whole blood culture assay. J Immunoassay. May-Aug;19(2-3):95-111.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1999) Differentiation between endotoxin and non-endotoxin pyrogens in human albumin solutions using an ex vivo whole blood culture assay. J Immunoassay. Feb-May; 20(1-2):79-89.

Poole S, Thorpe R, Meager A, Hubbard AJ, Gearing AJ (1988). Detection of pyrogen by cytokine release. Lancet 8577, 130.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, November 1989: 17-18.

Poole S, Musset MV (1989). The International Standard for Endotoxin: evaluation in an international collaborative study. J Biol Stand 17: 161-171.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' in vitro test for pyrogens. J Immunol Methods 274: 209-220.

Ray A, Redhead K, Selkirk S, Polle S (1991) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211-218.

Reddi, K., Nair, S.P. et al (1996). Surface-associated material from the bacterium Actinobacillus actinomycetemcomitans contains a peptide which, in contrast to lipopolysaccharide, directly stimulates fibroblast interleukin-6 gene transcription. Eur. J. Biochem. 236: 871-876.

Schindler S, Bristow A, Cartmell T, Hartung T and Fennrich S. (2003). Comparison of the reactivity of human and rabbit blood towards pyrogenic stimuli. ALTEX 20:59-63.

Schindler et al (2004) Cryopreservation of human whole blood for pyrogenicity testing. Journal of Immunological Methods 294, 89–100

Schins RP, van Hartingsveldt B, Borm PJ. Ex vivo cytokine release from whole blood. A routine method for health effect screening. Exp Toxicol Pathol. 1996 Nov; 48(6):494-6.

Spreitzer I, Fischer M, hartzsch K, Lüderitz-Püschel U and Montag T. (2000). Comparative Study of Rabbit Pyrogen Test and Human whole Blood Assay on Human Serum Albumin. ALTEX 19 (Suppl. 1):73-75

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

13.4 Original data

Include all available non-transformed original data for both the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method with established performance standards.

NOTE: The original data of the ELISA-plate reader were collected by S.Hoffman and ECVAM. These are available on the CD which goes with the BRD.

13.5 Performance standards

If appropriate performance standards for the proposed test method do not exist, performance standards for consideration by NICEATM and ICCVAM may be proposed. Examples of established performance standards can be located on the ICCVAM / NICEATM web site at http://iccvam.niehs.nih.gov.

BRD: WB/IL-6 March, 2006

APPENDIX A

Trial plan "Comparison And Validation Of Novel Pyrogen Tests Based On The Human Fever Reaction" Acronym: Human (e) Pyrogen Test

Detailed protocol WB/Il-6: *In vitro test for pyrogen/endotoxin using human whole blood 22 07 02 (electronic file name SOP WB IL-6)*

BRD: WB/IL-6 March, 2006

APPENDIX B

Not all of the publications cited in the BRD (see section 12) are included as hardcopies in Appendix B, e.g. publications on statistical methods are not given.

Remark: The same set of hardcopies was included into Appendix B of all of the 5 submitted BRDs. Therefore some of the publications in Appendix B might not be referenced in the current BRD nor included in the list of the publications in section 12. Three general publications were added, which are not cited in any BRD but might be useful as background information to the validation study: the ECVAM publications on validation (Balls et al, 1995; Curren et al, 1995; Hartung et al, 2004). Several publications were included, which either give more background information on the human fever reaction or report on specific studies using *in vitro* pyrogen tests.

List of hard copies

Andrade SS, Silveira RL, Schmidt CA, Junior LB, Dalmora SL. (2003) Comparative evaluation of the human whole blood and human peripheral blood monocyte tests for pyrogens. Int J Pharm. Oct 20;265(1-2):115-24.

Balls et al. (1995) Practical aspects of the validation of toxicity test procedures. ECVAM Workshop Report 5. ATLA 23, 129-147.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Carlin G, Viitanen E. (2003) In vitro pyrogenicity of a multivalent vaccine: Infanrix. PHARMEUROPA, 15(3), 418-423.

Curren et al (1995) The Role of prevalidation in the development, validation and acceptance of Alternative Methods. ECVAM Prevalidation Task Force Report 1. ATLA 23, 211-217.

De Groote, D., Zangerle, P.F., Gevaert, Y., Fassotte, M.F., Beguin, Y., Noizat-Pirenne, F., Pirenne, J., Gathy, R., Lopez, M., Dehart I., (1992). Direct stimulation of cytokines (IL-1 beta, TNF-alpha, IL-6, IL-2, IFN-gamma and GM-CSF) in whole blood. I. Comparison with isolated PBMC stimulation. Cytokine 4, 239.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Fennrich S, Wendel A and Hartung T. (1999). New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 16:146-149

Fennrich S, Fischer M, Hartung T, Lexa P, Montag-Lessing T, Sonntag H-G, Weigandt M und Wendel A. (1999). Detection of endotoxins and other pyrogens using human whole blood. Dev. Biol. Standards 101:131-139

Gaines Das et al (2004). Monocyte activation test for pro-inflammatory and pyrogenic contaminants of parenteral drugs: test design and data analysis. J. of Immunological Methods 288, 165-177.

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hartung et al (2004) A modular approach to the ECVAM principles on test validity. ATLA 32, 467-472.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005). Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Peterbauer A, Schindler S, et al (2005). International validation of novel pyrogen tests based on human monocytoid cells. J Immunol Methods 298: 161-173.

Jahnke M, Weigand, Sonntag H-G. (2000). Comparative testing for pyrogens in parenteral drugs using the human whole blood pyrogen test, the rabbit in vivo pyrogen test and the LAL test. European Journal of Parenteral Science 5(2):39-44

Nakagawa Y, Maeda H, Murai T. (2002) Evaluation of the in vitro pyrogen test system based on proinflammatory cytokine release from human monocytes: comparison with a human whole blood culture test system and with the rabbit pyrogen test. Clin Diagn Lab Immunol. May;9(3):588-97.

Pasare, C. and Medzhitov, R. (2003). Toll pathway-dependent blockade of CD4+CD25+ T cell-mediated suppression by dendritic cells. Science 299: 1033-1036.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1998) The detection of pyrogens in blood products using an ex vivo whole blood culture assay. J Immunoassay. May-Aug;19(2-3):95-111.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1999) Differentiation between endotoxin and non-endotoxin pyrogens in human albumin solutions using an ex vivo whole blood culture assay. J Immunoassay. Feb-May; 20(1-2):79-89.

Poole S, Thorpe R, Meager A, Hubbard AJ, Gearing AJ (1988). Detection of pyrogen by cytokine release. Lancet 8577, 130.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of

the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, November 1989: 17-18.

Poole S, Musset MV (1989). The International Standard for Endotoxin: evaluation in an international collaborative study. J Biol Stand 17: 161-171.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' in vitro test for pyrogens. J Immunol Methods 274: 209-220.

Ray A, Redhead K, Selkirk S, Polle S (1991) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211-218.

Reddi, K., Nair, S.P. et al (1996). Surface-associated material from the bacterium Actinobacillus actinomycetemcomitans contains a peptide which, in contrast to lipopolysaccharide, directly stimulates fibroblast interleukin-6 gene transcription. Eur. J. Biochem. 236: 871-876.

Schindler S, Bristow A, Cartmell T, Hartung T and Fennrich S. (2003). Comparison of the reactivity of human and rabbit blood towards pyrogenic stimuli. ALTEX 20:59-63.

Schindler et al (2004) Cryopreservation of human whole blood for pyrogenicity testing. Journal of Immunological Methods 294, 89–100

Schins RP, van Hartingsveldt B, Borm PJ. Ex vivo cytokine release from whole blood. A routine method for health effect screening. Exp Toxicol Pathol. 1996 Nov; 48(6):494-6.

Spreitzer I, Fischer M, hartzsch K, Lüderitz-Püschel U and Montag T. (2000). Comparative Study of Rabbit Pyrogen Test and Human whole Blood Assay on Human Serum Albumin. ALTEX 19 (Suppl. 1):73-75

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

BRD: WB/IL-6 March, 2006

APPENDIX C

List of abbreviations and definitions

Accuracy	The ability of a test system to provide a test result close to the accepted reference value for a defined property.
BET	The bacterial endotoxin test is used to detect or quantify endotoxins of gram-negative bacterial origin using amoebycte lysate from horseshoe crab (*Limulus polyphemus* or *Tachypleus tridentatus*
BRD	Background Review Document
CRYO WB/IL-1	Whole blood assay (using cryopreserved blood) with IL-1 as endpoint
CV	coefficient of variation
DL	Developing laboratory = laboratory which developed the method or the most experienced laboratory
ELC	Endotoxin limit concentration; maximum quantity of endotoxin allowed in given parenterals according to European Pharmacopoeia
Endotoxins	Endotoxins are a group of chemically similar cell-wall structures of Gram-negative bacteria, i.e. lipopolysaccharides
ELISA	Enzyme linked immunosorbent assay
EU/ml	European Units per ml
IL-1	interleukin 1
IL-6	interleukin 6
Intralaboratory reproducibility	A determination of the extent that qualified people within the same laboratory can independently and successfully replicate results using a specific protocol at different times.
Interlaboratory reproducibility	A measure of the extent to which different qualified laboratories, using the same protocol and testing the same substances, can produce qualitatively and quantitatively similar results. Inter-laboratory reproducibility is also referred to as between-laboratory reproducibility.
KN	University of Konstanz (Konstanz, Germany), developing laboratory WB/IL-1 and CRYO WB/IL-1
LPS	lipopolysaccharides
MM6	MONO MAC-6 cell line

MM6/IL-6	In vitro pyrogen test using MM6 cell line and IL-6 release as an endpoint
MVD	Maximum valid dilution; the MVD is the quotient of the ELC and the detection limit
NIBSC	National Institute for Biological Standards and Control (London, UK), developing laboratory for WB/IL-6
NL	naïve laboratory = laboratory with non or minor experience with the method
NPC	negative product control (clean, released lot of the nominated product under test)
Novartis	Novartis (Basel, Switzerland), developing laboratory PBMC/IL-6
OD	optical density
PBMC	Peripheral blood mononuclear cells
PBMC/IL-6	In vitro pyrogen test using fresh peripheral blood mononuclear cells and IL-6 release as endpoint
PBMC-CRYO/IL-6	In vitro pyrogen test using cryopreserved peripheral blood mononuclear cells and IL-6 release as endpoint
PEI	Paul-Ehrlich Institut (Langen, Germany), participating laboratory
PM	prediction model = is an explicit decision-making rule for converting the results of the in vitro method into a prediction of in vivo hazard
PPC	positive product control (product under test spiked with 0.5 EU/ml of WHO-LPS (code 94/580)
Prevalidation study	A prevalidation study is a small-scale inter-laboratory study, carried out to ensure that the protocol of a test method is sufficiently optimised and standardised for inclusion in a formal validation study. According to the ECVAM principles, the prevalidation study is divided into three phases: protocol refinement, protocol transfer and protocol performance (Curren et al, ATLA 23, 211-217).
Pyrogens	fever-causing materials
Pyrogens, endogenous	endogenous pyrogens are messenger substances released by blood cells reacting to pyrogenic materials; e.g. IL-1, IL-6, TNF-α, prostaglandin E_2
Pyrogens, exogenous	exogenous pyrogens derive from bacteria, viruses, fungi or from the host himself
Reliability	Measures of the extent to which a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is

	assessed by calculating intra- and interlaboratory reproducibility and intra-laboratory repeatability.
Relevance	Relevance of a test method describes whether it is meaningful and useful for a particular purpose. It is the extent to which the measurement result and uncertainty can accurately be interpreted as reflecting or predicting the biological effect of interest.
Repeatibility	Repeatability describes the closeness of agreement between test results obtained within a single laboratory when the procedure is performed independently under repeatability conditions, i.e. in a set of conditions including the same measurement procedure, same operator, same measuring system, same operating conditions and same location, and replicated measurements over a short period of time.
RIVM	National Institute of Public Health and the Environment (Bilthoven, The Netherlands), developing laboratory MM6/IL-6 method
Sensitivity	Sensitivity is the proportion of all positive/active substances that are correctly classified by a test method.
Specificity	Specificity is proportion of all negative/inactive substances that are correctly classified by a test method.
TMB	chromogenic substrate 3,3´,5,5´ -tetramethylbenzidine
TNF-α	tumour necrosis factor-α
USP	US Pharmacopoeia
Validation	Validation is the process by which the reliability and relevance of a procedure are established for a specific purpose
Validation study	A validation study is a large-scale interlaboratory study, designed to assess the reliability and relevance of an optimised method for a particular purpose
WB/IL-1	Whole blood assay (using fresh blood) with IL-1 release as endpoint
WB/IL-6	Whole blood assay (using fresh blood) with IL-6 release as endpoint
WHO	World Health Organization

Validation of Biomedical Testing Methods

In vitro test for pyrogen/endotoxin using human whole blood

22 07 02

Standard Operating Procedure

Only the responsible of the GLP/QA Unit is allowed to make copies of this document. Extra examples can be obtained at the GLP/QA Unit. Quality Documents are only valid if they are signed by the responsible of the GLP/QA Unit in and provided with a blue copy number	Copy number

Identity

Standard Operating procedure
In vitro *pyrogen test using human whole blood*

Version number:	220702
Applicable from:	22 07 02
Expired at:	

Drafted by:	Name	Stephen Poole
	Date	22 07 02
	Signature	
Reviewed by:	Name	
	Date	
	Signature	
Approved by:	Name	
	Date	
	Signature	
Issued by:	Name	
	Date	
	Signature	

*Owner/Trainer: Signature: Date:

PAGE OF CHANGES

Date of change/ Date of draft:	Version–number:	Changed page(s):	Summary of the change(s):	Changed by/Sign.:
22 07 02	220702		Sections 3, 8 and 10 redrafted	

TABLE OF CONTENTS ***Page No.***

1 INTRODUCTION....5
2 PURPOSE....5
3 SCOPE /LIMITATIONS....5
4 METHOD OUTLINE....5
5 DEFINITIONS ABBREVIATIONS....5
6 MATERIALS....5
7 METHODS....5
7.1. STEPS PRIOR TO BLOOD–CULTURE 5
7.1.1. Preparation of aliquots of the LPS (endotoxin) standard (STD)[at] 5
7.1.2. Preparation of aliquots of the IL–6 standard[at] 5
7.1.3. Coating of IL–6 ELISA plates.... 5
7.1.4. Preparation of samples for testing[at] 5
7.1.5. Collection of human blood 5
7.1.6. Storage of blood 5
7.1.7. Equilibration of reagents for cell culture.... 5
7.1.8. Preparation of the LPS standard curve[at].... 5
7.2. BLOOD CULTURE[at] 5
8 DATAANALYSIS....5
9 HEALTH SAFETY AND ENVIRONMENT....5
10 ANNEX....26
11 REFERENCES....5

1 INTRODUCTION

Parenteral pharmaceutical products must be shown to be free from pyrogenic (fever–inducing) contamination. While a pyrogen may in general be defined as any substance that causes fever, the pyrogens that almost invariably contaminate parenteral pharmaceuticals are bacterial endotoxins (lipopolysaccharides, LPS) from Gram–negative bacteria (Mascoli and Weary, 1979a, 1979b). There are two Pharmacopoeial tests for pyrogenic contamination: the rabbit pyrogen test and the Limulus amoebocyte lysate (LAL) test. The rabbit pyrogen test, which detects LPS and other pyrogens, involves measuring the rise in body temperature evoked in rabbits by the intravenous injection of a sterile solution of the substance to be examined. In contrast, the LAL test detects only LPS: it is described in Pharmacopoeias as the bacterial endotoxins test (BET). The principle of the LAL–test is that LPS causes extracellular coagulation of the blood (haemolymph) of the horseshoe crab, *Limulus polyphemus* (Levin & Bang, 1964). Although the LAL test is gradually superseding the rabbit pyrogen test, hundreds of thousands of rabbit pyrogen tests are still carried out each year around the world, largely on products which cannot, for one reason or another, be tested in the LAL test. While proving generally reliable, both the rabbit pyrogen test and LAL test have shortcomings. The rabbit pyrogen test uses experimental animals, is costly and is not quantitative. The LAL test gives false negatives with certain products, can overestimate the pyrogen content of other products and does not detect pyrogens other than bacterial endotoxin (LPS), such as Gram–positive exotoxins, viruses and fungi (Dinarello et al., 1984; Poole et al., 1988; Ray et al., 1990; Taktak et al., 1991; Fennrich et al., 1999).

The basis of the rabbit pyrogen test is the *in vivo* stimulation by exogenous pyrogens (usually LPS) of rabbit peripheral blood monocytes to produce the endogenous pyrogens that cause fever. The endogenous pyrogens are pyrogenic cytokines such as tumour necrosis factorα (TNFα), interleukin–1 (IL–1α and IL–1β, two separate gene products), IL–6 and IL–8 (Dinarello et al., 1999). In view of the shortcomings of the rabbit pyrogen test and the LAL test, *in vitro* pyrogen tests that utilise the exquisite sensitivity to exogenous pyrogen of monocytes have been proposed. In such tests, products are incubated with human peripheral blood monocytes (or mononuclear cells, PBMNC, or leukocytes) and the conditioned media assayed for pyrogenic cytokines (Duff & Atkins, 1982; Dinarello et al., 1984; Poole et al., 1988, 1989; Hansen and Christensen, 1990; Taktak et al., 1991; Bleeker et al., 1994).

The isolation of monocytes/leukocytes from whole blood is labour–intensive and time–consuming, technically sophisticated, requires expensive reagents and does not guarantee the isolation of cells in a non–activated state. Human whole blood produces cytokines in response to pyrogen/LPS (Desch et al., 1989; Finch–Arietta and Cochran, 1991; Hartung and Wendel, 1996) and *in vitro* pyrogen tests using human whole blood are being evaluated in a number of laboratories, particularly in Germany (Hartung and Wendel, 1996; Fennrich et al., 1999) and South Africa (Pool et al., 1998). The major differences between these whole blood *in vitro* pyrogen tests are the cytokine chosen as the readout (measured variable), the duration of incubation of the product (or LPS standard) with the blood, and the inclusion/omission of interferonγ as co–stimulus (priming agent). The preferred readout is IL–1β (Hartung and Wendel, 1996; Fennrich et al., 1999) or IL–6 (Pool et al., 1998), although TNFα also would appear to be suitable (Desch et al., 1989; Finch–Arietta and Cochran, 1991; Hartung and Wendel, 1996). The duration of incubation is determined to some extent by convenience but also by the choice of readout. A short

(4h) incubation of (diluted) blood with a very large dose of LPS (10 μg/ml, i.e. some 100000 IU/ml) stimulated the production of IL–1β (and TNFα) whereas a longer incubation (24h) of (diluted) blood was required for much smaller doses of LPS (1 pg/ml and above) to stimulate the production of IL–1β (and TNFα, Hartung and Wendel, 1996). A lengthy incubation (18 h), in the presence of interferonγ, was also favoured when IL–6 was the readout (Pool et al., 1998).

Our previous work with monocytes/monocytic cells pointed to the merits of IL–6 as the readout because IL–6, unlike IL–1β and TNFα, is secreted entirely into the cell–conditioned medium, in large quantities, permitting its complete estimation (Poole et al., 1989; Taktak et al., 1991).

2 PURPOSE

To develop an *in vitro* pyrogen test that will serve as a replacement for the rabbit pyrogen test.

3 SCOPE / LIMITATIONS

The method described below is for the evaluation of an in vitro whole blood/IL–6 release test. It is not a 'finalised' test system for the testing of medicinal products. The method may be applied only to preparations that have been validated with the method, i.e. shown not to interfere in the test system: see ANNEX.

This SOP is optimised for the detection of bacterial endotoxins. An alternative SOP, SOP.WBT.NEP.NIBSC.030702, is optimised for the detection of endotoxin and non-endotoxin pyrogens. SOP.WBT.NEP.NIBSC.030702 is to be used to test products suspected of contamination with non-endotoxin pyrogens.

4 METHOD OUTLINE

Freshly taken human whole blood is heparinised, diluted with saline and stimulated for 16–24h with standard endotoxin (LPS) and preparations under test. Following this stimulation, the concentration of IL–6 in the cell–conditioned medium is quantified using a specific ELISA (which is calibrated in terms of the appropriate international standard). The construction of a dose–response curve for endotoxin standard versus concentration of released IL–6 permits the estimation of the pyrogenic contamination of the preparations under test. The contamination is measured in endotoxin–equivalent units.

5 DEFINITIONS / ABBREVIATIONS

µg	microgram
µl	microlitre
Ab	antibody
BSA	bovine serum albumin
CO_2	Carbon dioxide
°C	degrees Celsius (Centigrade)
D–R	dose–response
E. coli	Escherichia coli
ELISA	Enzyme–linked immunosorbent assay
EP	European Pharmacopoeia
EU	endotoxin units
FDA	Food and Drug Administration (USA)
g	gram
h	hour
HIFCS	heat–inactivated (+56°C for 30 min) foetal calf serum
H_2O_2	hydrogen peroxide
H_2SO_4	sulphuric acid
IL	interleukin
IS	international standard
IU	international unit
l	litre
KOH	potassium hydroxide
LAL	limulus amoebocyte lysate
LPS	lipopolysaccharide
M	molar
MAb	monoclonal antibody
mg	milligram
min	minute
ml	millilitre
mM	millimolar
NaCl	sodium chloride
NaOH	sodium hydroxide
$NaHCO_3$	sodium hydrogen carbonate
NaH_2PO_4	sodium di–hydrogen phosphate
Na_2HPO_4	di–sodium hydrogen orthophosphate
No.	number
nm	nanometre
OD	optical density
PBMNC	peripheral blood mononuclear cells
PBS	Dulbecco's phosphate buffered saline
PC	Personal Computer
PF	pyrogen–free (items purchased as sterile and pyrogen–free or baked at 250°C for 30–60 min.
POD	horseradish peroxidase conjugate
R	endotoxin standard
rpm	rounds per minute
RSE	Reference Standard Endotoxin
RT	room temperature

TMB	Tetramethyl benzidine
S	test sample
STD	standard
UNK	unknown
USP	United States Pharmacopoeia
x g	x gravity

Institute for Health and Consumer Protection

6 MATERIALS

Sterile, pyrogen–free Dulbecco's phosphate buffered saline (Life Technologies)
HSA, 1% (a dilution in sterile PF saline of clinical grade HSA, 4.5%)
Polyoxyethylene–sorbitan monolaurate (TWEEN 20), cell culture grade, (Sigma, P–2287)
Hydrochloric acid, 0.1M, sterile filtered (Sigma, H–9892)
Sodium hydroxide (reagent grade)
1M H_2SO_4 (Merck)
Mouse monoclonal anti–IL–6 antibody from clone 16
Horseradish peroxidase conjugated sheep polyclonal anti–IL–6 antibody
3,3',5,5'–Tetramethyl benzidine (e.g. Fluka Cat. No. 87748)
Acetone (reagent grade)
Ethanol (reagent grade)
Phenol (e.g. Merck Cat. No. 100206)
Potassium hydroxide (reagent grade)
Sodium dihydrogen phosphate (e.g. Merck Cat. No. 106346)
Disodium hydrogen phosphate (e.g. Merck Cat. No. 106580)
Tris (hydroxymethyl) aminomethane (e.g. Fluka Cat. No. 93352)
Kathon MW/WT, Christ Chemie AG, Reinach, Switzerland
Albumin from bovine serum (e.g. Fluka Cat. No. 05480)
Citric acid monohydrate e.g. Fluka Cat. No. 27490)
Human AB serum (Sigma)
Trypan blue stain (Sigma)
USP Reference Standard Endotoxin [EC6 lot G], identical to the WHO international standards for bacterial endotoxin (LPS, vial code 94/580)
Fragmin (Dalteparin, 10000 IU/ml, Pharmacia)
Nunc–Immuno 96–well plate MaxiSorp (F96, Life Technologies™, Paisley, Scotland)
Falcon Microtest tissue culture plate, 96–well (353072, Beckton Dickinson Labware)
Falcon serological pipettes (5ml,10ml, 25ml, Beckton Dickinson Labware)
Centrifuge tubes (Falcon 2070 Blue Max™)
Polypropylene conical tubes (Falcon 2069 Blue Max™)
Eppendorf Biopur Tips 100ul & 1000ul (Eppendorf–Netheler–Hinz–Gmbh,Germany)
0.22 μm sterile filters (MilliPak 60, Millipore)
Eppendorf® volumetric pipettes
IL–6 from human lymphocytes (Boehringer Mannheim, Cat. No. 1299972)

All other consumables are purchased as sterile and pyrogen–free and other reagents are pro analysis grade.

Buffers for the NOVARTIS IL–6 ELISA:

Coating Buffer

Dissolve
5.0 g of sodium dihydrogen phosphate and
2.9 g of disodium hydrogen phosphate
in 400 ml of distilled water.

Use 1 N NaOH to adjust the pH to 7.5, and make up to 500 ml with distilled water.

Remains stable for 6 months at 2–8°C.

Blocking Buffer

Tris (hydroxymethyl)aminomethane	12.1 g
Dissolve in distilled water	400 ml
Add Kathon MW/WT	0.1 ml

Use 4 M HCl to adjust the pH to 7.5.

Albumin from bovine serum	5.0 g
Add distilled water to make up to 500 ml.	

Remains stable for 6 months at 2–8°C.

Stopping Solution

Distilled water	500 ml
H_2SO_4	26.6 ml

Wash Solution

Demineralised water	2000 ml
Tween–20	1 ml

Dilution Buffer

Prepare the dilution buffer as follows:

Tris(hydroxymethyl)aminomethane	2.1 g
Distilled water	400 ml
Kathon MW/WT,	0.1 ml
Phenol	0.5 g
Heat–inactivated (30 minutes at +56°C) foetal bovine serum	25 ml

Mix to dissolve the substances, then adjust the pH to 7.5 with 4 M HCl. Make up to 500 ml with distilled water.

Remains stable for at least 6 months at 2–8 C.
In the absence of the stabilizers Kathon and phenol the stability is only 1 day.

Institute for Health and Consumer Protection

TMB Solution

Prepare the TMB solution as follows:

3,3',5,5'Tetramethylbenzidine	240 mg
Reagent–grade acetone	5 ml

Dissolve, then add

Reagent–grade ethanol	45 ml
Perhydrol (30 % H_2O_2)	0.3 ml

Remains stable for at least 6 months at 15–25°C when sealed and protected from light.

Substrate Buffer

Reagent–grade citric acid monohydrate	6.3 g
Distilled water	800 ml

Mix to dissolve, then adjust the pH to 4.1 by adding 4 M KOH.
Make up to 1000 ml with distilled water and add 0.2 ml of Kathon MW/WT.

Remains stable for about 6 months at 15–25°C.
In the absence of the Kathon the stability is only 1 day.

7 METHODS

7.1. STEPS PRIOR TO BLOOD–CULTURE

Steps marked (*at*) are carried out in a Class 2 laminar flow sterile cabinet, using aseptic technique and reagents and consumables that are sterile and pyrogen–free.

7.1.1. Preparation of aliquots of the LPS (endotoxin) standard (STD)at

Make aliquots of the LPS STD: take a vial of the current IS for endotoxin (vial code 94/580, 10000 IU = EU/vial, infinite shelf life when stored at –20°C or below), open the vial, reconstitute the contents of the vial with 5 ml PFDW and vortex for 30 min. This gives a 2,000 IU/ml stock solution of LPS. The stock solution may be kept at 2–8°C for up to 14 days or frozen in aliquots immediately after reconstitution. The following primary standards (Reference Standard Endotoxins) are identical to the IS and may be substituted for it: EC6 (USP), Lot G (FDA), BRP3 (EP). Alternatively, a working standard (control standard endotoxin) calibrated against one of these primary standards may be substituted.

To prepare aliquots of the IS for endotoxin:

Aliquot 150 μl of LPS standard into labelled cryotubes (of 2 ml capacity), freeze them upright and store them in this frozen state at –20°C or below (shelf life = 12 months). Label the cryotubes with the following information:

Endotoxin STD 94/580
300 IU in 150 μl
date of reconstitution
initials of the operator

7.1.2. Preparation of aliquots of the IL–6 standardat

To make aliquots of the IL–6 standard, take a vial of the IS for IL–6 (ampoule code 89/548, 1 μg/100000 IU/ampoule), open the vial and reconstitute with 1 ml of PBS + 1% BSA (or HSA – not critical). This gives a 1 μg/ml stock solution of IL–6†.

(†: concentrations of IL–6 are expressed in pg/ml rather than IU/ml to avoid confusion with IU/ml of LPS.)

Aliquot 20 μl of the stock solution into labelled cryo–tubes (of 2 ml capacity), freeze them upright and store them in this frozen state at –20°C or below (shelf life = six months). Label the cryotubes with the following information:

IL–6 STD 89/548
20 ng in 20 μl
date of reconstitution
initials of the operator

(A working standard, previously calibrated against the IS may be substituted for the IS, Each new batch of working STD is to be calibrated against the IS.)

7.1.3. Coating of IL–6 ELISA plates

For the NOVARTIS IL–6 ELISA, dilute the coating anti–IL–6 antibody (Clone 16) with coating buffer to 2.5 μg/ml and swirl to mix, e.g. 1 mg of antibody in 400 ml of coating buffer. Add 200 μl to each well of a 96–well plate (Nunc–Immuno MaxiSorp F96). Stack the microtitre plates and allow to stand in the dark at 15–25°C for 16–24 h.

Aspirate and discard the coating solution. Wash the coated plate 3 times with demineralised water and tap out onto absorbent material, e.g. paper towel. Pipette 200 μl of blocking buffer into each well to block the residual protein–binding capacity of the coated plates. Seal the microtiter plates with adhesive film and store in a humidified atmosphere at 2–8°C (shelf life: two months).

7.1.4. Preparation of samples for testing[at]

Samples are tested at a dilution of 1 in 5, i.e. 50 μl of sample in a total culture volume of 250 μl. To test samples at dilutions greater than 1 in 5, pre–dilute samples before addition to the assay plate, e.g. to test a sample at a dilution of 1 in 10, pre–dilute the sample 1 in 2 with saline and add 50 μl of this diluted sample to the assay plate.

7.1.5. Collection of human blood

Qualification of blood donors: Blood donors are to describe themselves as being in good health, not suffering from any bacterial or viral infections (including colds and influenza), and to have been free from the symptoms of any such infection for a period of one week prior to the donation of blood. Blood donors are not to be taking non–steroidal anti–inflammatory drugs, immunosuppressants, glucocorticoids or any other drugs known to influence the production of cytokines. Also, See Section 8. (below) for the criteria for the rejection of data. Take blood donations from at least three donors because the preparation being examined is required to pass the test with blood donations from at least three different donors.

***Procedure*:** Using a 30 ml syringe and a 40 mm, 21 gauge hypodermic needle, draw 30 ml blood from the median cubital or cephalic vein of the left or right arm of a single donor.

Immediately remove the hypodermic needle and transfer the blood into a 50 ml sterile, pyrogen–free centrifuge tube (e.g. Falcon 2070 Blue Max) containing 10 IU in 10 μl heparin (Fragmin, 1000 IU/ml, prepared by diluting Fragmin, 10000 IU/ml, Pharmacia, 1/10 with saline) for each ml of blood to be collected, e.g. for a blood sample of 30 ml the tube will contain 300 IU/ml heparin, giving 10 IU heparin/ml of blood.

Screw the lid of the tube on tightly and invert slowly five times to ensure thorough mixing of the blood with the heparin. Do not vortex.

(N.B. The blood collection procedure is non-critical. Pyrogen-free heparin may be substituted for Fragmin and a proprietary blood collection device/system may be substituted provided that it is of a type shown to be pyrogen-free.)

Institute for Health and Consumer Protection

7.1.6. Storage of blood

Store the tube upright at room temperature and stimulate with LPS within 4 h of its collection.

7.1.7. Equilibration of reagents for cell culture

Bring an aliquot of the LPS standard, the samples for testing and a bottle of saline to room temperature.

7.1.8. Preparation of the LPS standard curve[at]

Prepare the LPS standard curve by making serial dilutions in saline of an aliquot of the stock solution of the current IS.

Label seven tubes, A – G. Add the volumes of saline to the tubes specified in table 1, below.

Add 1.35 ml saline to an aliquot (300 IU=EU in 150 µl) of the LPS standard and vortex to make 1.5 ml of a 200 IU/ml solution of LPS = Solution S.

Table 1. Preparation of the LPS standard curve:

Tube	LPS added to tube	Saline	[LPS] in tube	→ *[LPS] in well*
A	100 µl of Solution S = 20 IU	900 µl	20 IU/ml	*Not for culture*
B	200 µl of Solution A = 4 IU	800 µl	4 IU/ml	*0.8 IU/ml*
C	500 µl of Solution B = 2 IU	500 µl	2 IU/ml	*0.4 IU/ml*
D	500 µl of Solution C = 1 IU	500 µl	1 IU/ml	*0.2 IU/ml*
E	500 µl of Solution D = 0.5 IU	500 µl	0.5 IU/ml	*0.1 IU/ml*
F	500 µl of Solution E = 0.25 IU	500 µl	0.25 IU/ml	*0.05 IU/ml*
G	None	1 ml	0 IU/ml	*0 IU/ml*

Vortex each of Solutions A – G after its preparation and then use each solution for the preparation of the subsequent dilution.

(Multiples of the above volumes may be used to generate larger volumes of Solutions A – G and an LPS standard calibrated against the USP Reference Standard Endotoxin or the IS for endotoxin may be used.)

7.2. BLOOD CULTURE[at]

7.2.1. Add 50 µl of saline the wells of columns 1 – 10 as in Template 1, see below.

7.2.2. Gently mix the blood donation, using the Procedure below, immediately before aliquots of the blood are taken. Do not vortex the blood.

Procedure: Immediately after gently swirling the blood, pour a sample of about 5 ml into a smaller tube, e.g. a screw–top bottle of 7 ml capacity. Screw the lid on the tube and invert twice immediately before aliquots of the blood are taken. Replenish the blood in the smaller tube, as required, with further samples of blood, each taken immediately after gentle swirling of the blood donation.

Institute for Health and Consumer Protection

Institute for Health and Consumer Protection

7.2.3. Add 50 µl of blood to the wells of columns 1 – 10 as in Template 1, see below.

Using a pipette with a tip of wide diameter the blood is added by row to the wells of columns 1 – 10 in the following sequence: A, E, B, F, C, G, D, H (see Template 1, below). A repeating pipette may be used for these additions provided that the aliquots are added briskly to minimise the settling of cells.

Template 1: addition of blood to blood culture plate

	1	2	3	4	5	6	7	8	9	10	11	12
A	First row of ten wells (A1–A10) to which blood is added										Void	Void
B	Third row of ten wells (B1–B10) to which blood is added										Void	Void
C	Fifth row of ten wells (C1-C10) to which blood is added										Void	Void
D	Seventh row of ten wells (D1–D10) to which blood is added										Void	Void
E	Second row of ten wells (E1–E10) to which blood is added										Void	Void
F	Fourth row of ten wells (F1–F10) to which blood is added										Void	Void
G	Sixth row of ten wells (G1–G10) to which blood is added										Void	Void
H	Eighth row of ten wells (H1–H10) to which blood is added										Void	Void

7.2.4. Add 50 µl of LPS standards to wells as in Template 2, below.

Solution G into wells A3 – D3 (STD R0)
Solution F into wells A4 – D4 (STD R1)
Solution E into wells A5 – D5 (STD R2)
Solution D into wells A6 – D6 (STD R3)
Solution C into wells A7 – D7 (STD R4)
Solution B into wells A8 – D8 (STD R5)

(The above order of addition permits the same tip to be used for additions of all the standards.)

7.2.5. Add 50 µl of test samples S1 – S14 to wells, as in Template 2, see below.

7.2.6. Add 100 µl of saline to the wells of columns 1 – 10 as in Template 2, see below.

7.2.7. Gently mix the contents of the wells without cross–contaminating wells.

7.2.8. Incubate the cultures without vibration (to allow the cells to settle) at 37°C for 16– 24h in an atmosphere of 5% CO_2 in humidified air.

Template 2: addition of standards and samples to the blood culture plate

	1	2	3	4	5	6	7	8	9	10	11	12
A	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	void	void
B	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	void	void
C	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	void	void
D	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	void	void
E	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	void	void
F	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	void	void
G	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	void	void
H	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	void	void

Key:
S1 – S14 = test samples #1 – #14
R0 – R5 = Reference Standard Endotoxin, R0 = 0 IU/ml, R1 = 0.25 IU/ml, R2 = 0.5 IU/ml, R3 = 1 IU/ml, R4 = 2 IU/ml and R5 = 4 IU/ml (The final concentrations are: 0.05, 0.1, 0.2, 0,4 and 0.8 IU/ml).

7.3. IL–6 ELISA

Immunoreactive IL–6 in aliquots of the tissue culture fluid is quantified using a validated ELISA, in which the IS for IL–6 (89/548) or an IL–6 standard calibrated against the IS is used as the assay calibrant.

The NOVARTIS ELISA for IL–6 is described below. When it is not practicable to use the NOVARTIS ELISA, a different validated ELISA may be substituted, e.g. the MILIENIA IL-6 ELISA kit, CLB IL–6 ELISA kit, R&D IL-6 ELISA kit or an ELISA using matched pairs of anti–IL–6 MAbs purchased from R&D Systems or BioSource.

7.3.1. Equilibration of reagents

Bring an aliquot of the IL–6 standard and other assay reagents to room temperature before proceeding.

7.3.2. Preparation of IL–6 standard curve

About 30 min before the end of the tissue culture, prepare the IL–6 standard curve by making serial dilutions, of an aliquot of the stock solution of the current IS (or working STD for IL–6 calibrated against the IS). The diluent is PBS/1% BSA (or HSA – not critical).
(RPMI–C –see PBMNC SOP – may be substituted for PBS/1% BSA)
When using the IS for IL–6, label nine tubes H – P. Add the specified volumes of PBS/1% BSA to the tubes – see table 2 below.

Add 180 µl PBS/1% BSA to an aliquot (20 ng in 20 µl) of the IL–6 standard and vortex to make 200 µl of a 100 ng/ml solution of IL–6 = Solution G.

Table 2: Preparation of the IL–6 standard curve from an aliquot of the IS for IL–6

Tube	IL–6 added	PBS/1 % BSA	[IL–6] in tube	→ *[IL–6] in well*
H	100 µl of Solution G = 10 ng	900 µl	10 ng/ml	*Not for ELISA*
I	800 µl of Solution H = 8 ng	1.2 ml	4 ng/ml	*4000 pg/ml*
J	1 ml of Solution I = 4 ng	1 ml	2 ng/ml	*2000 pg/ml*
K	1 ml of Solution J = 2 ng	1 ml	1 ng/ml	*1000 pg/ml*
L	1 ml of Solution K = 1ng	1 ml	500 pg/ml	*500 pg/ml*
M	1 ml of Solution L = 500 pg	1 ml	250 pg/ml	*250 pg/ml*
N	1 ml of Solution M = 250 pg	1 ml	125 pg/ml	*125 pg/ml*
O	1 ml of solution N = 125 pg	1 ml	62.5 pg/ml	*62.5 pg/ml*
P	None	2 ml	0 pg/ml	*0 pg/ml*

Vortex each of Solutions H – P after its preparation and then use each solution for the preparation of the subsequent dilution.

(Multiples of the above volumes may be used to generate larger volumes of Solutions H – P and an IL–6 standard calibrated against the IS for IL–6 may be substituted for the IS.)

Store Solutions H – P at 2– 8°C until required.

7.3.3. Addition of samples

Immediately prior to adding standards and samples, empty the blocking buffer from the ELISA plate(s) and expel any remaining fluid by tapping the inverted plate onto absorbent material, e.g. paper towel.

At the end of the tissue culture incubation, carefully transfer 50 µl of supernatant from each of the wells of columns 1 – 10 of the tissue culture plate into the corresponding wells on the ELISA plate – see Template 1, above and Template 2, below. (The wells in columns 11 and 12 are for the cytokine standard curve – see below). The use of an eight channel pipette will facilitate the mixing and transfer. Change the pipette tips between each column transfer and ensure that cells are not aspirated by supporting the assay plate such that it slopes down at an angle of about 45° from column 1 to column 12.

Institute for Health and Consumer Protection

7.3.4. Addition of standards

Add 50 µl of IL–6 standards to the wells in columns 11 and 12, as shown in Template 3, below.

Solution P into wells A11 and A12 (0 pg/ml IL–6 STD)
Solution O into wells B11 and B12 (62.5 pg/ml IL–6 STD)
Solution N into wells C11 and C12 (125 pg/ml IL–6 STD)
Solution M into wells D11 and D12 (250 pg/ml IL–6 STD)
Solution L into wells E11 and E12 (500 pg/ml IL–6 STD)
Solution K into wells F11 and F12 (1000 pg/ml IL–6 STD)
Solution J into wells G11 and G12 (2000 pg/ml IL–6 STD)
Solution I into wells H11 and H12 (4000 pg/ml IL–6 STD)

(The above order of addition permits the same tip to be used for additions of all of the standards.)

Template 3: ELISA plate

	1	2	3	4	5	6	7	8	9	10	11	12
A	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	0	0
B	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	62.5	62.5
C	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	125	125
D	S1	S2	R0 0	R1 0.25	R2 0.5	R3 11	R4 2	R5 4	S3	S4	250	250
E	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	500	500
F	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	1000	1000
G	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	2000	2000
H	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	4000	4000

Key:
S1 – S14 = test samples #1 – #14
R0 – R5 = Reference Standard Endotoxin, R0 = 0 IU/ml, R1 = 0.25 IU/ml, R2 = 0.5 IU/ml, R3 = 1 IU/ml, R4 = 2 IU/ml and R5 = 4 IU/ml (The final concentrations are: 0.05, 0.1, 0.2, 0,4 and 0.8 IU/ml).Values 1 – 4000 in columns 11 and 12 are concentrations in pg/ml of the IS for IL–6 (ampoule code 94/580, 1 µg/ampoule, 1 pg = 0.1 IU).

7.3.5. Addition of (2nd) antibody–HRP conjugate (POD)

Add 200 µl of detection antibody POD (horseradish peroxidase conjugated to sheep anti–IL–6 antibodies: stable for at least 6 months at 2–8°C) pre–diluted with dilution buffer (usually 1/200 to 1/500, as determined in optimisation experiments) to each well, seal the plates with adhesive film, and allow to stand for 2–3 h at 20–25°C. (100 ml of diluted POD is sufficient for 4 ELISA plates.)

After incubation, wash plate three times with about 250 µl per well wash solution and then three times with demineralised water. Empty plate and expel any remaining fluid by tapping the inverted plate onto absorbent material, e.g. paper towel.

7.3.6. Addition of substrate solution and reading of optical densities

Prepare the substrate solution shortly before use. Transfer 90 ml of substrate buffer to a plastic bottle, add 4.5 ml of TMB solution and mix.

Pipette 200 µl of substrate solution into each well. After 10–15 minutes, stop the enzyme reaction by adding 50 µl/well of stopping solution. Wipe the back of the microtiter plates with a clean tissue, then measure the absorbance at 450 nm in an ELISA plate reader using a 540–590 nm corrective filter. Subtract the values of the measurement with the corrective filter from values measured with the 450 nm filter.

Institute for Health and Consumer Protection

8 DATA ANALYSIS AND PREDICTION MODELS

Preparatory testing

Preparatory tests are conducted to assure that the criteria for validity and precision for the standard curve are satisfied and that the solution of the preparation being examined (test article) does not interfere in the test. The test method is validated and a test for interfering factors repeated whenever there is any change in either the test method or the preparation being examined that is likely to influence the result of the test.

Validation of the test method

Using the standard endotoxin solution, prepare five endotoxin concentrations (R1–R5) to generate the standard curve (R0 = 0 EU/ml). Perform the test using four replicates of each standard endotoxin solution. The standard curve is to satisfy the criteria for linearity and range described in the *ICH HARMONISED TRIPARTITE GUIDELINE: VALIDATION OF ANALYTICAL PROCEDURES METHODOLOGY* (ICH guideline Q2B Validation of Analytical procedures: methodology, November 1996), together with a restriction on test precision (see below).

Assurance criteria for the standard curve

Reject the data from any test that does not satisfy all four criteria specified below.

(i) The (basal) release of IL–6 in the absence of added LPS, i.e. for the 0 IU/ml dose of LPS, is to be < 200 pg/ml IL–6 (20 IU IL–6/ml).

(The reason for this is as follows. Blood cultures in the absence of added endotoxin release amounts of (immunoreactive) TNFα and IL–1β close to or below the detection limits of the ELISAs used to detect them. In contrast, concentrations of immunoreactive IL–6 are detectable in these cultures. For healthy donors, concentrations of immunoreactive IL–6 are below 200 pg/ml (20 IU IL–6/ml, typically 50 pg/ml or less) but in donors who have recently recovered from minor viral or bacterial infections basal (i.e. unstimulated) concentrations of IL–6 can exceed 200 pg/ml. These findings accord with reports that plasma concentrations of immunoreactive IL–6 (unlike those of TNFα and IL–1β) are likely to be increased in individuals not in good health (Buck et al., 1994; Lin and Huang, 1998; Otto et al., 1999). Consequently, the choice of IL–6 as the readout (measured variable) enables the identification and 'screening out' (albeit retrospectively) of donors who are not in good health, even though they may feel well enough to serve as a blood donor for the test. Based on historic data obtained at NIBSC, a 'cut–off' value of 200 pg/ml IL–6 (20 IU IL–6/ml) for basal immunoreactive IL–6 release is deemed appropriate for the rejection of test data as having come from a test utilising blood from a donor not in good health.)

(ii) Compare the values for the four replicates each of 0 and 0.25 IU/ml endotoxin STD: the three smallest values are to be in the response to 0 IU/ml and the three largest values are to be in the response to 0.25 IU/ml (Wilcoxon rank-sum test, $p<0.05$).

(iii) Compare the values for the four replicates each of the endotoxin STDs 0.25 IU/ml and 0.5 IU/ml: the three smallest values are to be in the response to 0.25 IU/ml and the three largest values are to be in the response to 0.5 IU/ml (Wilcoxon rank-sum test, $p<0.05$).

(iv) Compare the values for the four replicates each of 0 and 0. 5 IU/ml endotoxin STD: the four smallest values are to be in the response to 0 IU/ml and the four largest values are to be in the response to 0.5 IU/ml (Wilcoxon rank-sum test, $p<0.025$).

Interference test

Spike an aliquot of the test solution of the preparation being examined, the test article (1), with an endotoxin concentration at or near the middle dose of the endotoxin standard curve. *[Alternatively, to facilitate comparisons with other test systems, a different spike, e.g. 0.5 IU/ml, may be used, provided that this is within one contiguous (adjacent) dilution of the middle dose of the endotoxin standard curve.]* Calculate the mean recovery of the added endotoxin by subtracting the mean endotoxin concentration in the test solution (if any) from that containing the added endotoxin. The target is 100% spike recovery. If necessary, test doubling dilutions of the test article, not exceeding the MVD (2), to identify the minimum dilution that gives spike recovery as close to 100% as can be achieved. Subsequent testing is to be performed at this dilution, which must give at least 50% spike recovery, or at one further two-fold dilution (3) of the test article, provided that the further dilution is within the MVD.

(1) the test article is to contain the lowest level of contamination with pyrogen/endotoxin that practicably can be achieved for that preparation.

(2) MVD = the maximum valid dilution of a preparation at which the endotoxin limit concentration (ELC) can be determined. MVD = ELC/the second smallest dose in the endotoxin STD curve.

(3) the further dilution is to mitigate variations in LOD, donors/passages, and the balance between interference by the test article and its pyrogenic content.)

(N.B. Testing at minimum dilution of the test article would be expected to increase the probability of detecting non-endotoxin pyrogens since, in general, these dilute out more rapidly than endotoxin. In this regard it should be noted that the interference testing described above, in common with the rest of this SOP, is optimised for the detection of bacterial endotoxins. A number of changes to this SOP would be required to optimise it for the detection of non-endotoxin pyrogens. Consequently a different protocol is used to test products suspected of contamination with non-endotoxin pyrogens.)

Quantitative estimation of endotoxin equivalents and uncertainty of measurement

Where quantitative estimates of endotoxin-equivalents are required, statistical analysis is carried out in Informatics to provide estimates of endotoxin concentrations for test samples in terms of the endotoxin standard curve using the methods of parallel line assay analysis (see for example, Finney, 1978).

The upper and lower OD response values for the endotoxin standard curve are determined graphically and the OD values for endotoxin standards and test samples are transformed to logit responses and analysed using the in house program WRANL (Gaines Das and Tydeman, 1982). The linearity mean square for the endotoxin standard curve (calculated by the program) provides a measure of the suitability of the graphically determined limits for the curve. The pyrogenic contamination in the solution of the preparation being examined (i.e. test article) is calculated, with confidence intervals, from a standard curve of the IS for endotoxin (which is calibrated in IU) and expressed in endotoxin-equivalents/ml. This value is compared with the endotoxin limit concentration (ELC) for the preparation. Where the ELC is not specified for a product, it is calculated as described in *10. ANNEX*. The preparation being examined complies with the test if the estimated mean endotoxin concentration of the preparation, after correction for dilution and concentration, is less than the endotoxin limit for the preparation. The preparation being examined is required to pass the test with blood donations from at least three different donors.

(Alternative analyses are permitted provided these are consistent with the relevant ICH guidelines.)

Limit test

This is a simplified test to detect whether or not a solution of the preparation being examined (test article), after correction for dilution and concentration, contains less than 0.5 IU/ml (the threshold dose in the rabbit pyrogen test). This is carried out as follows.

(a) Perform an interference test with an endotoxin spike of 0.5 IU/ml, as described above.

(b) Measure responses to 4 replicates each of 0, 0.25 and 0.5 IU/ml endotoxin and the test article.

(c) Reject the data from any test that does not satisfy all four criteria specified below.

(i) The (basal) release of IL–6 in the absence of added LPS, i.e. for the 0 IU/ml dose of LPS, is to be < 200 pg/ml IL–6 (20 IU IL–6/ml).

(ii) Compare the values for the four replicates each of 0 and 0.25 IU/ml endotoxin STD: the three smallest values are to be in the response to 0 IU/ml and the three largest values are to be in the response to 0.25 IU/ml (Wilcoxon rank-sum test, $p<0.05$).

(iii) Compare the values for the four replicates each of the endotoxin STDs 0.25 IU/ml and 0.5 IU/ml: the three smallest values are to be in the response to 0.25 IU/ml and the three largest values are to be in the response to 0.5 IU/ml (Wilcoxon rank-sum test, $p<0.05$).

(iv) Compare the values for the four replicates each of 0 and 0. 5 IU/ml endotoxin STD: the four smallest values are to be in the response to 0 IU/ml and the four largest values are to be in the response to 0.5 IU/ml (Wilcoxon rank-sum test, $p<0.25$).

(d) Compare the values for the four replicates each of the test article and the LPS STD 0.5 IU/ml: the four smallest values are to be in the response to the test article and the four largest values are to be in the response to 0.5 IU/ml (Wilcoxon rank-sum test, $p<0.25$). The preparation being examined is required to pass the test with blood donations from at least three different donors.

9 HEALTH SAFETY AND ENVIRONMENT

Human material

Human material should be treated as biologically hazardous and all work using human material is to be carried out according to the procedures specified in the NIBSC Safety Compendium.

Cultures of human material should be treated as biologically hazardous waste and disposed of according to the procedures specified in the NIBSC Safety Compendium.

Bacterial endotoxin is, as its name indicates, a toxic agent and should be handled with care.

Precautions: Cover open cuts before use. Do not get in eyes, on skin, on clothing. Avoid inhaling. Keep container closed.

First Aid: In case of contact, immediately flush eyes or skin with plenty of water for at least 15 minutes. If inhaled, remove to fresh air. If not breathing, give artificial respiration, preferably mouth–to–mouth. If breathing is difficult, give oxygen.

Effects of skin absorption can include fever, headache and hypotension.

Effects of inhalation can include fever, headache and hypotension.

Effects of ingestion – adverse effects are unlikely since ingested endotoxin is rapidly detoxified.

ELISA Substrate*: TMB** ***(Suspected mutagen, wear gloves when handling). **Store and use in accordance with manufacturer's instructions.**

10. ANNEX

LOD. Detection limit. The LOD is to take into account all dilution factors.

LOQ. Quantification limit. The LOQ is to take into account all dilution factors.

[N.B. The calculation of LOD and LOQ is addressed in the ICH HARMONISED TRIPARTITE GUIDELINE: VALIDATION OF ANALYTICAL PROCEDURES METHODOLOGY, ICH guideline Q2B Validation of Analytical procedures: methodology, November 1996. It should be noted that the methods described in the guideline, in common with other methods used to calculate LOD and LOQ, make assumptions about the distribution of data that may not apply.]

Calculation of ELC

ELC = Endotoxin limit concentration for the preparation being examined. Where an ELC is not specified, it is calculated as follows:
The sensitivity of rabbits to endotoxin is 5 IU/kg. So, for a product injected (i.e. tested) at 1 ml/kg, the detection limit is 5 IU endotoxin/ml/kg, giving an ELC of 5 IU endotoxin/ml, whereas for a product injected at 10 ml/kg, the detection limit is 5 IU endotoxin/10 ml/kg = 0.5 IU endotoxin/ml/kg, giving an ELC of 0.5 IU endotoxin/ml.

Maximum valid dilution

The MVD is the maximum allowable, i.e. 'valid', dilution of a preparation at which the endotoxin limit concentration (ELC) can be determined. MVD = ELC/the second smallest dose in the endotoxin STD curve.

Institute for Health and Consumer Protection

REFERENCES

Bleeker, W.K., de Groot, E.M., den Boer, P.J., Biessels, P.T., Aarden, L.A. and Bakker, J.C. (1994) Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol. 22, 835–40.

Buck, C., Bundschu, J., Gallati, H., Bartmann, P. and Pohlandt, F. (1994) Interleukin–6: a sensitive parameter for the early diagnosis of neonatal bacterial infection. Pediatrics 93, 54–8.

Desch, C.E., Kovach, N.L., Present, W., Broyles, C. and Harlan, J.M. (1989) Production of human tumor necrosis factor from whole blood ex vivo. Lymphokine Res. 8, 141–6.

Dinarello, C.A., Gatti, S. and Bartfai, T. (1999) Fever: links with an ancient receptor. Curr Biol. 9, R147–50.

Dinarello, C.A., O'Connor, J.V., LoPreste, G. and Swift, R.L. (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J. Clin. Microbiol. 20, 323–329.

Duff, G.W. and Atkins, E. (1982) The detection of endotoxin by *in vitro* production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52, 323–332.

Fennrich, S., Fischer, M., Hartung, T., Lexa, P., Montag–Lessing, T., Sonntag, H.G., Weigandt, M. and Wendel. (1999) A Detection of endotoxins and other pyrogens using human whole blood. Dev Biol Stand. 101, 131–9.

Finch–Arietta, M.B. and Cochran, F.R. (1991) Cytokine production in whole blood ex vivo. Agents Actions 34, 49–52.

Finney DJ: Statistical Method in Biological Assay, Third Edition. London, Charles Griffin and Company Ltd, 1978.

Gaines Das RE and Tydeman MS. (1982) Iterative weighted regression analysis of logit responses: A computer program for analysis of bioassays and immunoassays. Computer Programs in Biomedicine 15; 13–22.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15, 425–33.

Hartung, T. and Wendel, A. (1996) Detection of pyrogens using Human Whole Blood. In vitro Toxicology 9, 353–359.

Levin, J. and Bang, F.B. (1964) A description of cellular coagulation in the Limulus. Bull. John Hopkins Hosp. 155, 337–345.

Institute for Health and Consumer Protection

Lin. S.J. and Huang J.L. (1998) Circulating interleukin (IL)–1 beta, IL–6 and tumor necrosis factor–alpha in children with febrile infection–a comparison with C–reactive protein. Asian Pac J Allergy Immunol 1998. 16, 105–9.

Mascoli, C.C. and Weary, M.E. (1979a) Limulus amebocyte lysate (LAL) test for detecting pyrogens in parenteral injectable products and medical devices: advantages to manufacturers and regulatory officials. J Parenter Drug Assoc. 33, 81–95.

Mascoli, C.C. and Weary, M.E. (1979b) Applications and advantages of the Limulus amebocyte lysate (LAL) pyrogen test for parenteral injectable products. Prog Clin Biol Res. 29, 387–402.

Otto, G., Braconier, J., Andreasson, A. and Svanborg, C. (1999) Interleukin–6 and disease severity in patients with bacteremic and nonbacteremic febrile urinary tract infection. J Infect Dis. 179, 172–9.

Pool, E.J., Johaar, G., James, S., Petersen, I. and Bouic, P. (1998) The detection of pyrogens in blood products using an ex vivo whole blood culture assay. J Immunoassay 19, 95–111.

Poole, S. (1989) 'Monocyte test' for pyrogen. Pharmeuropa, 1(5), 238.

Poole, S., Selkirk, S., Rafferty, B., Meager, A., Thorpe, R. and Gearing, A. (1989*a*) Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa, 1 (*Special Issue*),1989, 17–18.

Poole, S., Thorpe, R., Meager, A., Hubbard, A.R. and Gearing, A.J.H. (1988) Detection of pyrogen by cytokine release. Lancet 8577, 130.

Ray, A., Redhead, K., Selkirk, S. and Poole, S. (1990) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211–218.

Schins, R.P.F., van Hartingsveldt, B., and Borm and P.J.A. (1996) *Ex vivo* cytokine release from whole blood. Exp. Toxic Pathol. 48, 494–496.

Taktak, Y.S., Selkirk, S., Bristow, A.F., Carpenter, A., Ball, C., Rafferty, B. and Poole, S. (1991) Assay of pyrogens by interleukin–6 release from monocytic cell lines. J Pharm Pharmacol 43, 578–582.

Institute for Health and Consumer Protection

Appendix A4

The Human Peripheral Blood Mononuclear Cell (PBMC)/IL-6 *In Vitro* Pyrogen Test

ECVAM Background Review Document (March 2006) A-337
ECVAM Standard Operating Procedure (July 2004) A-411
ECVAM Standard Operating Procedure for the Catch-Up Validation Phase (October 2002) A-433

[This Page Intentionally Left Blank]

BRD: PBMC/IL-6 March, 2006

IN VITRO PYROGEN TEST USING HUMAN PBMCs (PBMC/IL-6)

BRD: PBMC/IL-6 March, 2006

Contents

1 RATIONALE FOR THE PROPOSED TEST METHOD 4

1.1 INTRODUCTION 4
1.2 REGULATORY RATIONALE AND APPLICABILITY 5
1.3 SCIENTIFIC BASIS FOR THE PROPOSED TEST METHOD 6

2 TEST METHOD PROTOCOL COMPONENTS 8

2.1 OVERVIEW OF TEST METHOD 8
2.2 RATIONAL FOR SELECTED TEST COMPONENTS 9
2.3 BASIS FOR SELECTION OF THIS TEST METHOD 13
2.4 PROPRIETARY COMPONENTS 13
2.5 REPLICATES 13
2.6 MODIFICATIONS APPLIED AFTER VALIDATION 14
2.7 DIFFERENCES WITH SIMILAR TEST METHODS 14

3 SUBSTANCES USED FOR VALIDATION 15

3.1 SELECTION OF SUBSTANCES USED 15
3.2 NUMBER OF SUBSTANCES 15
3.3 DESCRIPTION OF SUBSTANCES USED 16
3.4 SAMPLE CODING PROCEDURE 17
3.5 RECOMMENDED REFERENCE CHEMICALS 17

4 *IN VIVO* REFERENCE DATA ON ACCURACY 18

4.1 TEST PROTOCOL *IN VIVO* REFERENCE TEST METHOD. 18
4.2 ACCURACY 18
4.3 ORIGINAL RECORDS 19
4.4 QUALITY OF DATA 19
4.5 TOXICOLOGY 20
4.6 BACKGROUND ON ASSAY PERFORMANCE 20

5 TEST METHOD DATA AND RESULTS 21

5.1 TEST METHOD PROTOCOL 21
5.2 ACCURACY AND RELIABILITY 23
5.3 STATISTICS 29
5.4 TABULATED RESULTS 33
5.5 CODING OF DATA 35
5.6 CIRCUMSTANCES 35
5.7 OTHER DATA AVAILABLE 35

6 TEST METHOD ACCURACY 36

6.1 ACCURACY 36
6.2 CONCORDANCY TO *IN VIVO* REFERENCE METHOD 38
6.3 COMPARISON WITH REFERENCE METHODS 38
6.4 STRENGTH AND LIMITATIONS 39
6.5 DATA INTERPRETATION 39
6.6 COMPARISON TO OTHER METHODS 39

7 TEST METHOD RELIABILITY (REPEATABILITY/REPRODUCIBILITY) 40

7.1 SELECTION OF SUBSTANCES 40
7.2 RESULTS 40
7.3 HISTORICAL DATA 45
7.4 COMPARISON TO OTHER METHODS 45

8 TEST METHOD DATA QUALITY 46

8.1 CONFORMITY 46
8.2 AUDITS 46
8.3 DEVIATIONS 46
8.4 RAW DATA 46

9 OTHER SCIENTIFIC REPORTS AND REVIEWS 47

9.1 SUMMARY 47
9.2 DISCUSSION 48
9.3 RESULTS OF SIMILAR VALIDATED METHOD 48

10 ANIMAL WELFARE CONSIDERATIONS (REFINEMENT, REDUCTION, AND REPLACEMENT) 51

10.1 DIMINISH ANIMAL USE 51
10.2 CONTINUATION OF ANIMAL USE 51

11 PRACTICAL CONSIDERATIONS 53

11.1 TRANSFERABILITY 53
11.2 TRAINING 53
11.3 COST CONSIDERATIONS 54
11.4 TIME CONSIDERATIONS 54

12 REFERENCES 55

13 CATCH -UP VALIDATION: HUMAN PBMC/IL-6 IN VITRO PYROGEN TEST WITH CRYOPRESERVED MONOCYTOID CELLS. 57

13.1 RATIONALE 57
13.2 TEST METHOD PROTOCOL COMPONENTS 57
13.3 SUBSTANCES USED FOR VALIDATION 57
13.4 TEST METHOD ACCURACY 57
13.5 TEST METHOD RELIABILITY (REPRODUCIBILITY) 60
13.6 CONCLUSION 61

14 SUPPORTING MATERIALS (APPENDICES) 62

14.1 STANDARD OPERATING PROCEDURE (SOP) OF THE PROPOSED METHOD 62
14.2 STANDARD OPERATING PROCEDURE (SOP) OF THE REFERENCE METHOD 62
14.3 PUBLICATIONS 62
14.4 ORIGINAL DATA 65
14.5 PERFORMANCE STANDARDS 65

Appendix A – Method protocol(s) and trial plan(s)

Appendix B – Hardcopies of relevant publications

Appendix C – List of abbreviations and definitions

BRD: PBMC/IL-6 March, 2006

1 Rationale for the Proposed Test Method

1.1 Introduction

1.1.1. *Describe the historical background for the proposed test method, from original concept to present. This should include the rationale for its development, an overview of prior development and validation activities, and, if applicable, the extent to which the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards.*

Pyrogens, a chemically heterogeneous group of hyperthermia- or fever-inducing compounds, derive from bacteria, viruses, fungi or from the host himself, and provoke an immune response by producing endogenous pyrogens such as prostaglandins and the pro-inflammatory cytokines interleukin-1 (IL-1), interleukin-6 (IL-6) and tumour necrosis factor-α (TNF-α) (Dinarello, 1999). Depending on the type and amount of pyrogen challenge and the sensitivity of an individual, even life-threatening shock-like conditions can be provoked. To assure quality and safety of any pharmaceutical product for parenteral application in humans, pyrogen testing is therefore imperative.

Depending on the medicinal product, one of two animal-based pyrogen tests is currently prescribed by the respective Pharmacopoeias, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET). For the rabbit pyrogen test, sterile test substances are injected intravenously to rabbits and any rise in body temperature is assessed. This *in vivo* test detects various pyrogens but not alone the fact that large numbers of animals are required to identify a few batches of pyrogen-containing samples argues against its use. In the past two decades, the declared intention to refine, reduce and replace animal testing, has lowered rabbit pyrogen testing by 80% by allowing to use the BET as an *in vitro* alternative pyrogen test for certain medicinal products (Cooper et al, 1971).

Bacterial endotoxin, comprising largely lipopolysaccharide (LPS) from the cell wall of Gram-negative bacteria that stimulates monocytes/macrophages via interaction with CD14 and toll-like receptor 4 (TLR4) (Beutler and Rietschel, 2003), is the pyrogen of major concern to the pharmaceutical industry due to its ubiquitous sources, its stability and its high pyrogenicity. With the BET, endotoxin is detected by its capacity to coagulate the amoebocyte lysate from the haemolymph of the American horseshoe crab, *Limulus polyphemus*, a principle recognised some 40 years ago (Levin and Bang, 1964). In the US, *Limulus* crabs are generally released into nature after drawing about 20% up to 50% of their blood and therefore most of these animals survive. However, the procedure still causes mortality of about 30,000 horseshoe crabs per year, which adds to even more efficient threats of the horseshoe crab population like its use as bait for fisheries, habitat loss and pollution (http://www.horseshoecrab.org/).

As with the rabbit test the general problem of translation of the test results to the human fever reaction persists. Moreover, although being highly sensitive, the failure of the BET to detect non-endotoxin pyrogens as well as its susceptibility to interference by e.g. high protein or lipid levels of test substances or by glucans impedes full replacement of the rabbit pyrogen test. Hence, hundreds-of-thousands rabbits per year are still used for pyrogen testing.

A test system that combines the high sensitivity and *in vitro* performance of the BET test with the wide range of pyrogens detectable by the rabbit pyrogen test is therefore

required in order to close the current testing gap for pyrogens and to avoid animal-based tests. With this intention and due to improved understanding of the human fever reaction (Dinarello, 1999), test systems based on *in vitro* activation of human monocytes were developed. First efforts date back about 20 years, when peripheral blood mononuclear cells (PBMC) were used to detect endotoxin by monitoring the release of pyrogenic cytokines (Duff and Atkins, 1982; Dinarello et al, 1984). Meanwhile, a number of different test systems, using either whole blood, peripheral blood mononuclear cells (PBMCs) or the monocytoid cell lines MONO MAC 6 (MM6) or THP-1 as a source for human monocytes and various read-outs were established (Poole et al., 1988; Ziegler et al, 1988; Tsuchiya et al, 1980; Hartung and Wendel, 1996; Hartung et al, 2001; Poole et al, 2003). These test systems were validated with the aim of developing a tool for formal inclusion into Pharmacopoeias, an important basis for implementing novel alternative pyrogen tests for product-specific validation.

1.1.2 *Summarize and provide the results of any peer review conducted to date and summarize any ongoing or planned reviews.*
When the validation study was commenced, the proposed test method had not been subjected to any other peer review process. In the meantime, the PBMC/IL-6 test developed by Novartis and Baxter Healthcare has been subjected to a rigorous peer-review by the US FDA and approved as an end-product release test (New Drug Application Number 16-267/S-037 approved on April 24, 2002).
All of the five methods are currently under peer review of the ECVAM Scientific Advisory Committee.

1.1.3 *Clearly indicate any confidential information associated with the test method; however, the inclusion of confidential information is discouraged.*
This document does not contain any confidential information.

1.2 Regulatory rationale and applicability

1.2.1 *Describe the current regulatory testing requirement(s) for which the proposed test method is applicable.*
To assure quality and safety of pharmaceutical products for parenteral application in humans, pyrogen testing is imperative. Depending on the drug, one of two pyrogen tests is currently prescribed by the European Pharmacopoeia, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET), and other national and international guidelines.

1.2.2 *Describe the intended regulatory use(s) (e.g., screen, substitute, replacement, or adjunct) of the proposed test method and how it will be used to substitute, replace, or complement any existing regulatory testing requirement(s).*
Dependent on the product and the presence of relevant clinical data on unexpected pyrogenicity of clinical lots, the proposed test method may be an alternative method for pyrogen testing, thus substituting the rabbit pyrogen test or in rare cases the BET. In certain cases, the proposed test method may function as a supplementary test method to assess compliance to the licensing dossier.
In case the proposed test method is an alternative for pyrogenicity testing, a thorough cross-validation between the proposed test method and the original method for the

specific medicinal product is warranted. In case the proposed test method is an adjunctive test to screen for (unexpected) pyrogenic lots, alert and alarm limits may be established based on consistency of production lots or (preferably) based on actual clinical data.

1.2.3 *Where applicable, discuss the similarities and differences in the endpoint measured in the proposed test method and the currently used in vivo reference test method and, if appropriate, between the proposed test method and a comparable validated test method with established performance standards.*
The current *in vivo* method (rabbit test), as described in the pharmacopoeia, and the proposed *in vitro* test method each determine very different end-points, though the biochemical origins of the response are similar.
The *in vivo* method determines the total rise in body temperature (fever induction) of the animals subjected to the medicinal product, as a result of pyrogens present in the product.
The proposed PBMC/IL-6 test is an *in vitro* model for the fever response mechanism. It determines the release of cytokines by monocytoid cells into the culture medium upon the interaction of pyrogens and specific receptors on the monocytoid cells. It is these cytokines that trigger the fever response *in vivo*.
Main difference between the *in vivo* and *in vitro* method are that the latter is quantitative and uses cells of human origin, thus better reflecting the physiological situation.

1.2.4 *Describe how the proposed test method fits into the overall strategy of hazard or safety assessment. If a component of a tiered assessment process, indicate the weight that should be applied relative to other measures.*
The proposed PBMC/IL-6 test method may be applied for those medicinal products for which the rabbit test is the only or most reliable method for pyrogenicity testing, since a) the medicinal product is not compatible with the BET or b) the medicinal product contains pyrogens other than Gram-negative endotoxin.
Limit concentrations for pyrogens are established based on consistency lots or actual clinical data or, in the case of endotoxin the endotoxin limit concentration (ELC) as defined for many medicinal products.

1.3 Scientific basis for the proposed test method

1.3.1 *Describe the purpose and mechanistic basis of the proposed test method.*
The proposed *in vitro* method is intended to determine the presence of pyrogens in medicinal products for parenteral use. The proposed test method is an *in vitro* model of the human fever response. It determines the release of cytokines upon the interaction of pyrogens and specific Toll-like receptors on the monocytoid cells (Beutler and Rietschel, 2003). It is these cytokines that trigger the fever response *in vivo*.

1.3.2 *Describe what is known and not known about the similarities and differences of modes and mechanisms of action in the proposed test method as compared to the species of interest (e.g., humans for human health-related toxicity testing).*
An important feature of the proposed test method is that it is based upon the use of monocytoid cells of human origin. It therefore by definition resembles more closely the actual response of humans. The two other test methods make use of either crustaceans (BET) or rabbits, both species more or less distinct from the human species. The response

of humans, horseshoe crabs and rabbits toward Gram-negative endotoxin has been studied extensively and the methods appear equivalent for this particular pyrogen (Cooper et al 1971; Greisman and Hornick, 1969). However, there are documented cases of medicinal products and specified pyrogenic substances that yield false-positive or false-negative results in either test method. Since the proposed test method is based on human cells, it may therefore predict more accurately the pyrogenicity of such substances in humans.

1.3.3 *Describe the intended range of substances amenable to the proposed test method and/or the limits of the proposed test method according to chemical class or physicochemical factors.*

The proposed test method is intended for the assessment of pyrogens in all parenteral medicinal products for human use, chemical or biological and including raw materials, bulk ingredients and excipients. Use of the proposed test method in testing environmental samples or medicinal products is suggested and may be feasible, but substantiating data are as yet limited or absent.

2 Test Method Protocol Components

2.1 Overview of test method

Provide an overview of how the proposed test method is conducted. If appropriate, this would include the extent to which the protocol for the proposed test method adheres to established performance standards.

A highly detailed protocol describing the proposed test method (*Detailed protocol PBMC/IL-6: "In Vitro Pyrogen Test Using PBMC (SP+PB var. Novartis) 03 10 02; electronic file name: SOP-PBMC IL 6*) is attached in Appendix A of this Background Review Document (BRD).

The PBMC/IL-6 test method is an assay for the detection of pyrogenic contamination. The test protocol itself can be divided into the following two parts:

1. Incubation of the sample with peripheral blood mononuclear cells (PBMCs).
2. An enzyme linked immunoassay (ELISA) for the measurement of IL-6.

Ad 1.
Human whole blood is collected by venipuncture into heparinized tubes for blood sampling. The blood is diluted with PBS and the PBMCs are obtained by density gradient centrifugation (Lymphoprep). PBMCs from 4 donors are incubated overnight (16-24 hours) together with saline and the sample of interest in sterile and pyrogen-free cell culture plate.

Ad 2.
After mixture of the samples, the concentrations of the cytokine IL-6 in the PBMC-conditioned medium are quantified using a ELISA specific for IL-6 as follows.
Aliquots are added to the wells of a microtiterplate coated with IL-6 specific monoclonal antibodies. An enzyme-conjugated polyclonal antibody against IL-6 is added. During a subsequent incubation, a sandwich complex consisting of two antibodies and the IL-6 is formed. Unbound material is removed in between incubation steps by a washing of the ELISA-plate.
A chromogenic substrate (3,3´,5,5´ -tetramethylbenzidine [TMB]) reactive with the enzyme label is added. Color development is terminated by adding a stopping solution. The resulting color, read at the appropriate wavelength, is directly related to the IL-6 concentration.
The IL-6 ELISA used throughout this study is an in-house ELISA, developed by Novartis, in which the IL-6 calibrant is calibrated against the International Standard (IS) for IL-6 (WHO code: 89/548).

The WHO-LPS standard (code 94/580, E.coli O113:H10:K-), was used throughout the validation. This standard is identical to USP Reference Standard Endotoxin (EC6).

There are several possibilities to estimate the pyrogenic contamination of the preparations under test: 1) A quantitative estimation can be achieved by the construction of a dose-response curve for endotoxin standard (e.g. 5.0, 2.5, 1.0, 0.5 and 0.25 EU/ml) versus

Optical Density (OD)-value of the IL-6 ELISA. The contamination of the preparations is expressed in Endotoxin Unit Equivalents. 2) A qualitative test can be achieved by the inclusion of an endotoxin threshold control (e.g. one fixed dilution of the standard curve) which allows for the classification in positive and negative samples (i.e. pyrogenic and non-pyrogenic samples). 3) A qualitative test can also be achieved by inclusion of an appropriate positive product control.
A detailed description of analysis methods used during the validation of the test-method can be found in section 5 of the current BRD.

2.2 Rational for selected test components

Provide a detailed description and rationale, if appropriate, for the following aspects of the proposed test method:
2.2.1 Materials, equipment, and supplies needed.
The materials, equipment and supplies used for the PBMC/IL-6 test method are laboratory items that will be already available in a routine QC laboratory. There is no need for sophisticated or dedicated laboratory equipment throughout the test.
For all steps in the procedure, excluding the ELISA procedure, the materials (e.g. tips, containers, solutions) which will be in close contact with samples and blood cells need to be sterile and pyrogen free. The materials, equipment and supplies are specified in the method protocol (Appendix A). It should be noted that equivalent devices may also be used and it is the user's responsibility to validate the equivalence.

Materials for part 1: PBMC-Incubation
Equipment
- Incubator (37°C, 5% CO_2 humidified air)
- Class 2 laminar flow sterile cabinet
- Centrifuge (suitable for 50 ml centrifuge tubes)
- Vortex

Consumables
- Lymphoprep (Nycomed, Oslo, Norway)
- Fragmin
- 30 ml syringe with luer lock adapter
- 19 mm, 21 gauge butterfly needle (or 40 mm, 21 gauge butterfly needle)
- 5, 10, 25 ml serological pipettes.
- Pipettes suitable for 50 or 100 µl
- 96-wells tissue culture plates
- WHO-LPS standard

Materials for part 2: ELISA procedure
Equipment
- Multichannel pipettor.
- Pipettes suitable for 50, 100 or 1000 µl.
- Microplate mixer
- Microplate washer
- Microplate reader capable of readings at the appropriate wavelength

- A software package facilitating data generation, analysis, reporting, and quality control

Consumables

- Graduated cylinder and plastic storage containers
- 96-wells microtiter plates
- Mouse monoclonal anti-IL-6 antibody from clone 16 (Novartis)
- Horseradish peroxydase conjugated sheep polyclonal anti-Il-6 antibody.
- Human Interleukin-6 standard
- Coating buffer, blocking buffer, dilution buffer, stopping solution and wash solution as detailed in the test method protocol.

The IL-6 ELISA used is an in-house assay developed in the Novartis laboratory and uses the WHO IL-6 international standard. Any commercially available IL-6 ELISA kit using the same standard or a standard calibrated versus it may be used (if validated for this in-vitro pyrogen test). Including the appropriate positive and negative controls in each run ensures the reliability and accuracy of the PBMC/IL-6 test method. As a positive control a specified amount of the Endotoxin Standard is used. The assay should be considered acceptable only if the criteria described in the method protocol are met. Also the criteria for allowed variability of replicates within an assay have to be met. The IL-6 standard curve is an additional control of the performance of the assay.

2.2.2 Dose-selection procedures, including the need for any dose range-finding studies or acute toxicity data prior to conducting a study, if applicable.

For every kind of test compound the interference with PBMC and the Il-6 ELISA is determined. For this purpose a preliminary "dose finding" test is conducted to establish a suitable dilution for every new test compound. For the validation study (as described in section 4 of this BRD), the tested products were diluted according to their known ELC, which was usually far beyond interfering concentrations. The ELC of the tested products or drugs were calculated according to the European Pharmacopoeia. If no endotoxin limit is defined it can be estimated by dividing 350 EU by the maximum hourly dose (example: the maximum hourly dose is 100 mg/patient, then the estimated endotoxin limit is 350/100=3.5EU/mg).

2.2.3 Endpoint(s) measured.

The proposed test method is an *in vitro* model of the fever response mechanism. It determines the release of interleukin-6 (IL-6) by monocytoid cells isolated from human blood. IL-6 is released into the culture medium upon the interaction of pyrogens and specific receptors on the monocytoid cells. The measured endpoint IL-6 is one of the cytokines that trigger the fever response *in vivo*.

2.2.4 Duration of exposure.

The PBMCs are exposed to possible pyrogenic components in samples at 37°C for 16-24 hours in an atmosphere of 5% CO_2 in humidified air. This conditioned mixture containing endogenous pyrogens released by the cells, is subsequently assayed in the IL-6 ELISA.

2.2.5 Known limits of use.
The PBMC/IL-6 test as described in the test method protocol is not a finalized test system for the testing of all medicinal products. The method may be applied only to preparations that have been validated with this method, i.e. shown not to interfere with the blood and the Il-6 readout system at a specified dilution of the preparation. A paragraph describing the interference testing is included in the method protocol (Appendix A). However, at this moment there are no medicinal products known that cannot be tested with the method.

2.2.6 Nature of the response assessed.
The proposed test method is an *in vitro* model of the fever response mechanism. Upon the interaction of exogenous pyrogens and specific receptors on the monocytoid cells endogenous pyrogens (e.g. interleukins, TNF-□ and prostaglandins) are produced. In the body the fever response is triggered by these endogenous pyrogens. Immunoreactive IL-6, the measured endpoint for the current method, is one of these endogenous pyrogens.

2.2.7 *Appropriate vehicle, positive, and negative controls and the basis for their selection*.
Throughout the development and validation phase the test compounds are diluted in 0.9% (w/v) clinical saline. This 0.9% clinical saline is considered an appropriate vehicle as no interference with active substances of a drug is to be expected.
In addition the test includes several controls.
A negative control: 0.9% clinical saline (sodium chloride)
A positive control: WHO-LPS 94/580, 0.5 EU/ml in clinical saline.
A negative product control (NPC): clean, released batch for each drug.
A positive product control (PPC): test item spike with WHO-LPS (code 94/580) at 0.5 EU/ml.
The positive and negative controls are the same in every assay and are needed to establish the sensitivity of the test system. In addition, a product-based set of controls is used to reveal product related interference.

2.2.8 Acceptable range of vehicle, positive and negative control responses and the basis for the acceptable ranges.
A PBMC/IL-6 assay is considered acceptable for further analysis if the following criteria are met: The ELISA is valid if the OD of the blank control is below 0.15 and the mathematical function (quadratic model) of the IL-6 standard curve produces an r^2 >0.95. The reactions (in terms of OD) on the endotoxin concentrations give a sigmoidal ascending dose response.

Exclusion criteria for blood donors (low and high responders):
Blood donors are considered low responders if their mean OD value for the endotoxin reference standard concentration 1 EU/ml is below the mean OD value for 1000 pg/ml of IL-6. Blood donors are considered high responder if the mean OD value for the negative control above the mean OD value at 500 pg/ml of IL-6

If the test samples show an irregular response (e.g. high SD), the results of the blank and the standard endotoxin concentrations derived from the donor are checked. The donor is excluded if these results are inconsistent with the ordinarily expected results.

As regards the substances to be tested, for products with an established ELC, specified in EU/ml, the product is diluted to its maximum valid dilution (MVD). The negative product control should be negative at the MVD. The response to the positive product control should be between 50% and 200% of the response to the positive control, indicating a possible pyrogenicity can be detected using these conditions.

However, for the results described throughout this BRD the data were accepted and analyzed according to the procedures described in section 5.3 "Statistics". This procedure was chosen as it allowed for a harmonized analysis of comparable data which were obtained with different *in vitro* pyrogen tests (i.e. MM6/IL-6, WBT/IL-1, WBT/IL-6).

2.2.9 Nature of the data to be collected and the methods used for data collection.
The raw data collected are the read-outs (absorbance) of the IL-6 ELISA, measured by an automated laboratory ELISA-plate reader. The wavelength is dependent on the chromogenic substrate applied. In the current study TMB is used and the ELISA-plate is read at a wavelength of 450 nm using a 540 nm to 590 nm corrective filter. The values of the measurement with the corrective filter is subtracted from values measured with the 450 nm filter.

2.2.10 Type of media in which data are stored.
Data are stored in electronic files (windows98 compatible software) and as hard copy.

2.2.11 Measures of variability.
As part of the development of the PBMC/IL-6 test method the intralaboratory repeatability was assessed by independent and identical replicated measurement of the different concentrations of WHO-LPS. Furthermore, the limit of detection and its dependence from known but uncontrollable variables such us operator and donor were investigated. These variables and the inherent variation of biological systems make up to the total variation of the method.

2.2.12 Statistical or non-statistical methods used to analyze the resulting data, including methods to analyze for a dose-response relationship. Justify and describe the method(s) employed.
The first phase of the study was performed with PBMCs from 1 individual donor. The large scale blinded studies with pharmaceutical drugs were run with PBMCs from 4 individual donors in parallel. A standard, using the International Standard for Endotoxin (calibrated in EU) is included, ranging from 0.063 EU/ml up to 1.0 EU/ml.

2.2.13 Decision criteria and the basis for the prediction model used to classify a test chemical (e.g., positive, negative, or equivocal), as appropriate.
A prediction model (PM) was developed in order to classify substances as "pyrogenic for humans" or "non-pyrogenic for humans". To be able to define a dichotome result in the

alternative pyrogen test, a threshold pyrogen value of 0.5 EU/ml was chosen. This threshold value was based on historical data with rabbits (described in section 4.1). The suitability of the PM was assessed by testing substances which were artificially contaminated with endotoxin. (substances are described in section 3.2 and 3.3)
The statistical approach, including quality criteria, is detailed in section 5.3

2.2.14 Information and data that will be included in the study report and availability of standard forms for data collection and submission.
Raw data were collected using a standard form. These were submitted to the quality department of ECVAM.

2.3 Basis for selection of this test method

Explain the basis for selection of the test method system. If an animal model is being used, this should include the rationale for selecting the species, strain or stock, sex, acceptable age range, diet, and other applicable parameters.
In view of the shortcomings of the rabbit pyrogen test and the BET, *in vitro* pyrogen tests that utilize the exquisite sensitivity to exogenous pyrogen of monocytoid cells have been proposed. In such tests, products are incubated with human cell and the conditioned media assayed for pyrogenic cytokines (Duff & Atkins, 1982; Dinarello et al., 1984; Poole et al., 1988; Poole et al, 1989; Hansen & Christensen, 1990; Taktak et al., 1991; Bleeker et al., 1994).
The human PBMC assay was developed as a real *in vitro* alternative to the rabbit pyrogen test. The basic idea was to mimic the fever reaction in humans. In general, the detection of exogenous pyrogens (e.g. endotoxin) by monocytoid cells causes them to release endogenous pyrogens like IL-1β, IL-6 and TNFα. These cytokines affect the thermal regulation centre in the brain and increase the body temperature by changing its set point. In the past several test methods have been developed that use the sensitivity of human peripheral blood monocytes to exogenous pyrogens. In an attempt to increase the sensitivity of these tests the monocytes/leukocytes were isolated from whole blood, as has been done for the PBMC/IL-6 test. An overview of relevant literature can be found in section 9 of this BRD. Interleukin IL-6 is chosen as the readout because IL-6, unlike IL-1 and TNF, is secreted entirely into the cell-conditioned medium in large quantities, thereby permitting its complete estimation.

2.4 Proprietary components

If the test method employs proprietary components, describe what procedures are used to ensure their integrity (in terms of reliability and accuracy) from "lot-to-lot" and over time. Also describe procedures that the user may employ to verify the integrity of the proprietary components.
The test method does not employ proprietary components.

2.5 Replicates

Describe the basis for the number of replicate and repeat experiments; provide the rationale if experiments are not replicated or repeated.
All experiments with medicinal drugs are run using individual PBMCs, isolated from 4 individual donors.

During a prevalidation phase, the intralaboratory reproducibility as well as the interlaboratory reproducibility of the PBMC/IL-6 test method was established by applying repeated experiments (see section 7). As the test method reliability (repeatability /reproducibility) was shown to be satisfactory, it was feasible to establish the accuracy using pharmaceutical substances (detailed in table 3.3.1) by one test performed by three participating laboratories (see section 6).

2.6 Modifications applied after validation

Discuss the basis for any modifications to the proposed test method protocol that were made based on results from validation studies.
During the first part of the validation, which was restricted to the testing of LPS-spiked saline, only one donor was used. To reduce the variation the test was performed with PBMCs of 4 donors assayed separately.

The test can easily be adjusted to a quantitative assay as described in the detailed method protocol. However, the assay has now been validated as a qualitative assay, by means of the PM.

2.7 Differences with similar test methods

If applicable, discuss any differences between the protocol for the proposed test method and that for a comparable validated test method with established performance standards.
Not applicable.

3 Substances Used for Validation

3.1 Selection of substances used

Describe the rationale for the chemicals or products selected for use in the validation process. Include information on the suitability of the substances selected for testing, indicating any chemicals that were found to be unsuitable.

Selected test items were medicinal products available on the market. Released clinical batches were considered clean, i.e. containing no detectable pyrogens. To test the specificity, sensitivity and the reproducibility of the proposed test method, the products were spiked with pyrogen. For the present studies endotoxin (LPS) was selected as the model pyrogen, since it is well defined, standardized and readily available.

For the sensitivity and specificity the test items were assessed at their maximal valid dilution (MVD). The MVD is the quotient of the ELC and the detection limit. The European Pharmacopoeia prescribes for various types of parenterals the amount of endotoxin that is maximally allowed in a medicinal product, i.e. the ELC, taking into consideration the dose, the route of administration and the dosing regimen of the product. The aim of the study was to discriminate between negative and positive samples. Based on the selected pyrogen threshold value (see 4.1) and the PM applied, positive samples were defined as containing 0.5 EU/ml or more. Hence, to determine the MVD, the value of 0.5 EU/ml was defined as the detection limit.
Test items were assessed as such (negative product control), spiked with endotoxin at 0.5 IU/ml (positive product control) and after spiking with endotoxin at 5 levels (blinded samples). In addition a negative control (saline) and positive control (0.5 EU/ml in saline) were included to establish assay validity.

For reproducibility, the test items were tested, at a predefined dilution above the MVD, independently in 3 different laboratories, 3 times each. Based on the selected pyrogen threshold value (see 4.1) and the PM applied, positive samples were defined as containing 0.5 EU/ml or more. The test items were tested after spiking with endotoxin at four levels. For no other reasons but practical ones, i.e. availability of test materials, different test items were selected for this part of the validation study.

It was determined earlier whether candidate test items interfered with the outcome of the proposed test method. Interference was considered when the response of endotoxin in the diluted test item was below 50% or above 200% of the response of endotoxin in saline (spike-recovery). It was shown that none of the test items interfered with the assay at the selected dilutions (data not shown).

3.2 Number of substances

Discuss the rationale for the number of substances that were tested.

A total of 13 test items were selected for the validation study (see 3.3): 10 test items for determining sensitivity and specificity (table 3.3.1), 3 different test items for determining reproducibility (table 3.3.2). Test items and their spikes were appropriately blinded by ECVAM before distribution to the participating testing facilities.

For sensitivity and specificity, each test item was tested after spiking at its individual MVD. Hence they each came with their own specific set of 5 endotoxin spike solutions: 0.0, 0.25, 0.5, 0.5 and 1.0 EU/ml. Simple logistics limited the amount of test items for this part of the validation study to 10. Since test items were assessed with 5 different endotoxin levels at 3 independent test facilities, this yielded a total of 150 data points, biometrically considered to be sufficient for further analysis.

For reproducibility each test item was spiked at 4 different levels (0.0, 0.0, 0.5 and 1.0 EU/ml) and tested at specified dilutions, 3 times at 3 laboratories.

3.3 Description of substances used

Table 3.3.1: Test items (parenteral drugs) used for determining sensitivity and specificity

Drug	code	Source	Agent	Indication	MVD (-fold)
Glucose 5% (w/v)	GL	Eifel	Glucose	nutrition	70
Ethanol 13% (w/w)	ET	B.Braun	Ethanol	diluent	35
MCP®	ME	Hexal	Metoclopramid	antiemetic	350
Orasthin®	OR	Aventis	Oxytocin	initiation of delivery	700
Binotal®	BI	Aventis	Ampicillin	antibiotic	140
Fenistil®	FE	Novartis	Dimetindenmaleat	antiallergic	175
Sostril®	SO	GlaxoSmithKline	Ranitidine	antiacidic	140
Beloc®	BE	Astra Zeneca	Metoprolol tartrate	heart dysfunction	140
Drug A*	LO	-	0.9% NaCl	-	35
Drug B*	MO	-	0.9% NaCl	-	70

*Drugs A and B were included as saline controls using notional ELCs.
Negative control: 0.9% clinical stock saline solution.
Positive control: WHO-LPS 94/580 (*E. coli* 0113:H10:K-), 0.5 EU/ml in clinical stock saline.

Table 3.3.2: Test items (parenteral drugs) used for determining reproducibility.

Drug	Source	Agent	Indication
Gelafundin®	Braun melsungen	Gelatin	Transfusion
Jonosteril ®	Fresenius	Electrolytes	Infusion
Haemate ®	Aventis	Factor VIII	Hemophilia

Negative control: 0.9% clinical stock saline solution.
Positive control: WHO-LPS 94/580 (*E. coli* 0113:H10:K-), 0.5 EU/ml in clinical stock saline.

3.4 Sample coding procedure

Describe the coding procedures used in the validation studies.
All test items are registered medicinal products and were obtained from a pharmaceutical supplier. Test items and endotoxin spiking samples were prepared, blinded where appropriate and coded under GLP by personnel from ECVAM, Italy. These were then taken over by the Paul-Ehrlich Institute, Germany, for testing-allocation and shipment to each of the appropriate test facilities participating in the study.
For the sensitivity and specificity part of this study, test items and their respective spikes (5 per test item) were all blinded. For reproducibility, only the spikes (4) were blinded, the test items were not.

3.5 Recommended reference chemicals

For proposed test methods that are mechanistically and functionally similar to a validated test method with established performance standards, discuss the extent to which the recommended reference chemicals were tested in the proposed test method. In situations where a listed reference chemical was unavailable, the criteria used to select a replacement chemical should be described. To the extent possible, when compared to the original reference chemical, the replacement chemical should be from the same chemical/product class and produce similar effects in the in vivo reference test method. In addition, if applicable, the replacement chemical should have been tested in the mechanistically and functionally similar validated test method. If applicable, the rationale for adding additional chemicals and the adequacy of data from the in vivo reference test method or the species of interest should be provided.
The reference pyrogen material used was the international endotoxin standard WHO-LPS 94/580 (*E. coli* 0113:H10:K-). Where appropriate, the material was diluted in clinical saline solution (0.9%(w/v) sodium chloride). The saline was also used as negative control (blank).

4 *In vivo* Reference Data on Accuracy

4.1 Test protocol *in vivo* reference test method.

Provide a clear description of the protocol(s) used to generate data from the in vivo reference test method. If a specific guideline has been followed, it should be provided. Any deviations should be indicated, including the rationale for the deviation.
For ethical reasons, no rabbit pyrogen tests were performed for this study. However, Dr. U. Lüderitz-Püchel, Paul-Ehrlich Institute, Germany, kindly provided historical data, accumulated over several years, from 171 rabbits (Chinchilla Bastards). The respective Pharmacopoeia's do not prescribe a rabbit strain for the *in vivo* pyrogen test, but Chinchilla rabbits are reported as a relatively sensitive strain for pyrogen testing.

The rabbits were injected with endotoxin and their rise in body temperature over the next 180 minutes was recorded (figure 4.1.1). From these data it was established that 50% of the rabbits got fever when treated with 5 EU/Kg (Hoffmann et al, 2005a). Fever in rabbits is defined as a rise in body temperature over 0.55°C. On the basis of these historical animal data and corrected for the maximal volume allowed in rabbits, i.e. 10 mls per animal, a pyrogen threshold value of 0.5 EU/ml was defined for the PM in the proposed test method.

4.2 Accuracy

Provide the in vivo reference test method data used to assess the accuracy of the proposed test method. Individual human and/or animal reference test data, if available, should be provided. Provide the source of the reference data, including the literature citation for published data, or the laboratory study director and year generated for unpublished data.
As mentioned, no animal studies were done for ethical reasons. However, a theoretical assumption on specificity and sensitivity can be made (Hoffmann et al, 2005a). Taking the prevalence of the different final concentrations of LPS in the 5 spikes into account (1.0 EU/ml: 20%; 0.5 EU/ml: 40%; 0.25 EU/ml: 20% and 0.0 EU/ml: 20%) and calculating probabilities of misclassification using the chosen study design and defined threshold of pyrogenicity, i.e. 0.5 EU/ml, the theoretical sensitivity of the rabbit pyrogen test is 57.9% and the theoretical specificity of the rabbit pyrogen test is 88.3%.

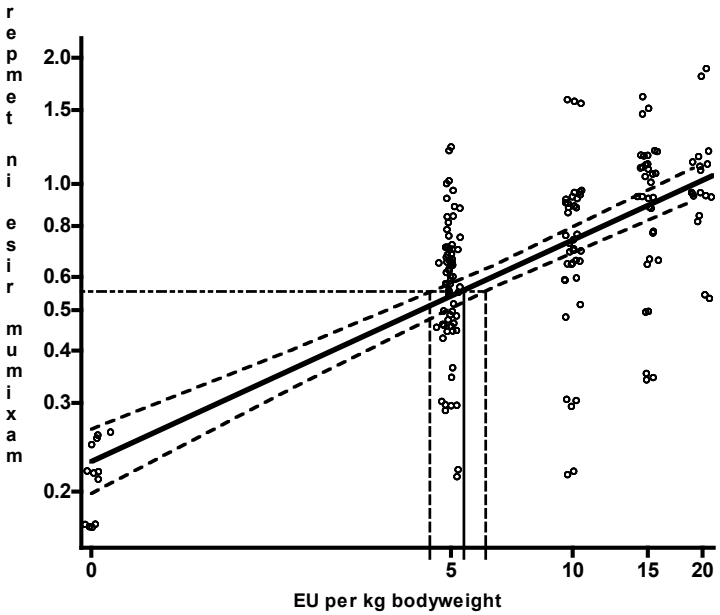


Figure 4.1.1 Dose-temperature of standard endotoxin applied to Chinchilla Bastards (n=171). Rabbits were treated with 1 ml saline containing 0, 5, 10, 15 and 20 EU of *E. coli* LPS (WHO-LPS 94/580 (E.coli O113:H10:K)) and their body temperature was measured over 180 min. Linear regression analysis was performed after logarithmic transformation of the data. Data are shown as dots to which a jitter-effect was applied in order to be able to distinguish congruent data. The full line depicts the linear regression whereas the dashed lines represent the 95%-confidence bounds. Furthermore, a horizontal line for a 0.55°C raise of temperature is added which is often defined as the rabbit threshold for fever. At the interception point of this line and the regression line 50% of the rabbits are to be expected to develop fever.

4.3 Original records

If not included in the submission, indicate if original records are available for the in vivo reference test method data.

The recognition of pyrogenic substances as bacterial by-products and the identification of a variety of pyrogenic agents enabled the development of a proper test to demonstrate non-pyrogenicity of the pharmaceutical product. As early as the 1920s, studies were done to select the most appropriate animal model. Results indicated that most mammals had a pyrogenic response, but only a few, including rabbits, dogs, cats, monkeys and horses showed a response similar to that in humans. For practical reasons, other species but rabbits and dogs were considered not practical. In 1942, Co Tui & Schrift described that rabbits are less thermo-stable as compared to dogs. Hence, rabbits are more suited for the purpose of testing for the absence of pyrogens, since a negative result is more significant.

4.4 Quality of data

Indicate the quality of the in vivo reference test method data, including the extent of GLP compliance and any use of coded chemicals.

All procedures employed in the study are GLP-concordant and quality assured by ECVAM's quality assurance officers.

4.5 Toxicology

Discuss the availability and use of relevant toxicity information from the species of interest (e.g., human studies and reported toxicity from accidental or occupational exposure for human health-related toxicity testing).

Over time, a number of studies were done to correlate the rabbit test to pyrogenic reactions in humans. A conclusive study by Greisman and Hornick, published in 1969, who compared three purified endotoxin preparations (*Salmonella typhosa*, *E. Coli* and *Pseudomonas*) in New Zealand rabbits and in male volunteers, showed that the induction of a threshold pyrogenic response, on a weight basis, was similar to rabbit and man. At higher doses, rabbits respond less severe as compared to man.

4.6 Background on assay performance

Discuss what is known or not known about the accuracy and reliability of the in vivo reference test method.

As mentioned, no animal studies were done for ethical reasons. However, a theoretical assumption on specificity and sensitivity can be made (Hoffmann et al, 2005a). Taking the prevalence of the different final concentrations of LPS in the 5 spikes into account (1.0 EU/ml: 20%; 0.5 EU/ml: 40%; 0.25 EU/ml: 20% and 0.0 EU/ml: 20%) and calculating probabilities of misclassification using the chosen study design and defined threshold of pyrogenicity, i.e. 0.5 EU/ml, the theoretical sensitivity of the rabbit pyrogen test is 57.9% and the theoretical specificity of the rabbit pyrogen test is 88.3%.

5 Test Method Data and Results

5.1 Test method protocol

Describe the proposed test method protocol used to generate each submitted set of data. Any differences from the proposed test method protocol should be described, and a rationale or explanation for the difference provided. Any protocol modifications made during the development process and their impact should be clearly stated for each data set.

The protocol for the PBMC/IL-6 test is provided Appendix A of this BRD. It includes the precise step-by-step description of the test method, including the listing of all the necessary reagents and laboratory procedures for generating data. For two steps during validation a part of the method protocol was adapted to contain a detailed description of the dilution of the samples and the spiking with WHO-LPS. The relevant part of the protocol is detailed in this section as well. The validity criteria and the detailed statistical analysis described in section 5.3 of this BRD were applied to analyse the data produced during validation.

To assess the reliability of the test method a series of experiments were conducted in the developing laboratory (DL). As a start, only blanks (saline, 0 EU/ml) and spikes of WHO-LPS in saline were tested. These experiments are summarised in table 5.1.1.

Table 5.1.1: summary of experiments with WHO-LPS in saline.

Experiment	Spikes (EU/ml) in saline	n (per spike)	Repetitions of experiment	N
1A	0; 0.25; 0.5	20	1	60
1B	0; 0.063; 0.125; 0.25; 0.5	12	1	60
2A	0; 0.5	8	3	48
2B	0; 0.063; 0.125; 0.25; 0.5	8	3	120
2C	0; 0.125; 0.25; 0.5	8	8	256

The collected data were used to answer questions regarding the nature of the distribution, the variance and its behaviour over the range of response in replicated measurements under identical conditions. In addition intralaboratory reproducibility was assessed by the maintenance of minimum criteria, e.g. a detection limit below an a-priori chosen positive control or a dose-dependent standard curve (table 5.1.1, experiment 1b). With the data of this experiment an assessment of the limit of detection of the test can be done by calculating the smallest spike, which can be discriminated from the blank.

The PBMC/IL-6 method was transferred from the DL to two other laboratories (denoted as naive laboratory 1 [NL1] and naive laboratory 2 [NL2]). A large scale dose response experiment was performed by all three laboratories. For this study 6 or 7 concentrations were tested in a dose response curve (typically 0; 0.125; 0.25; 0.5; 1; 2 EU/ml, at least 8 replicates) and all laboratories had to meet the validity criteria as laid down in the method protocol before the studies with medicinal substances were conducted.

The **intra- and interlaboratory reproducibility** was assessed by testing 3 different medicinal substances, Gelafundin, Jonosteril and Haemate (described in table 3.3.2). Test items and their spikes were appropriately blinded. Test items were tested, at a predefined dilution above the MVD, independently in 3 laboratories, 3 times each. Test items were tested after spiking with WHO-LPS at four different levels, the spikes were blinded and coded by QA ECVAM. In addition a negative control (saline) and positive control (0.5 EU/ml) in saline were included to establish assay validity.
Although this part of the study was designed for assessment of reproducibility, a preliminary estimate of the accuracy could be derived from the data. Applying the PM to the results and evaluating concordance in a two-by-two contingency table assessed accuracy.

To assess **accuracy** of the proposed test method 10 substances (listed in table 3.3.1), were spiked with five different concentrations of the WHO-LPS (one of which is negative). To permit the application of the chosen PM, each drug was diluted to its individual MVD, which has been calculated using the ELC to that drug (listed in table 3.3.1.) Each substance had their own specific set of endotoxin spike solution, ensuring that after spiking the undiluted drug, it contained 0.0; 0.25; 0.5; 0.5 and 1.0 EU/ml respectively when tested at its MVD. A detailed description of the sample preparation was supplied to the three independent laboratories (shown in table 5.1.2). To put more weight to the result of this part of the validation, the spikes were blinded and coded by QA ECVAM. The raw data of the PBMC/IL-6 assay are shown in paragraph 5.2. Accuracy was assessed by applying the PM to the results and evaluating the concordance in a two by two table. As intralaboratory reproducibility was (successfully) shown in previous experiments, only interlaboratory reproducibility was assessed in this phase.

Table 5.1.2: Sample preparation for the testing of 10 substances spiked with 5 different concentrations of WHO-LPS.

unblinded			*blinded*				
dilution of drug up to **MVD** ↓			spiking of **undiluted** drug: 0.5 ml each				
diluted drug	**NPC**	**PPC**	+ 23.3 µl of Spike 1	+ 23.3 µl of Spike 2	+ 23.3 µl of Spike 3	+ 23.3 µl of Spike 4	+ 23.3 µl of Spike 5
0.5 ml	+ 25 µl saline	+ 25 µl PPC-LPS-spike *					
		(final conc. = 50 pg/ml)	**dilution** to **MVD** ↓				
	test	test	test	test	test	test	test

** PPC-LPS-spike contains 1050 pg/ml = 21fold 50 pg/ml*

NPC = Negative Product Control, **PPC** = Positive Product Control, **MVD** = Maximal Valid Dilution

5.2 Accuracy and reliability

Provide all data obtained to evaluate the accuracy and reliability of the proposed test method. This should include copies of original data from individual animals and/or individual samples, as well as derived data. The laboratory's summary judgment regarding the outcome of each test should be provided. The submission should include data (and explanations) from all studies, whether successful or not.
See figures 5.2.1, 5.2.2, 5.2.3, 5.2.4, 5.2.5 (A, B and C), 5.2.6 and 5.2.7 (A and B).

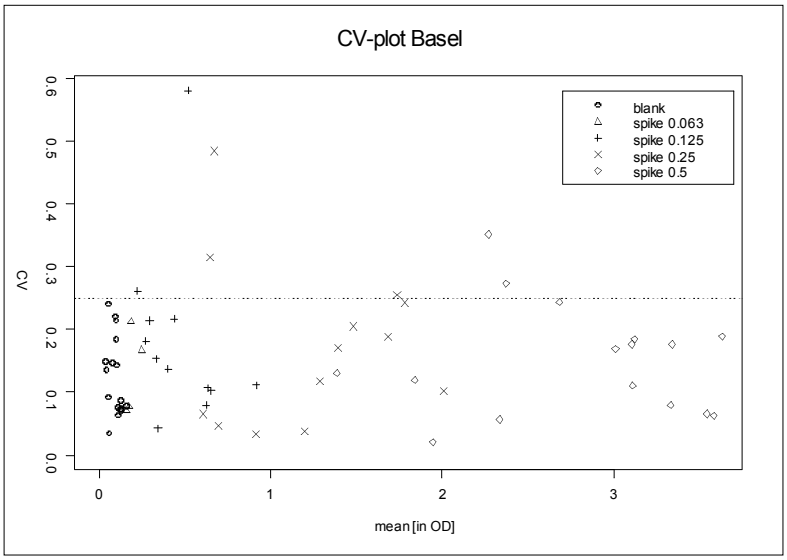


Figure. 5.2.1: Coefficient of variation (CV) of WHO-LPS spikes (4 replicates) relative to the mean OD (readout of the IL-6 ELISA).

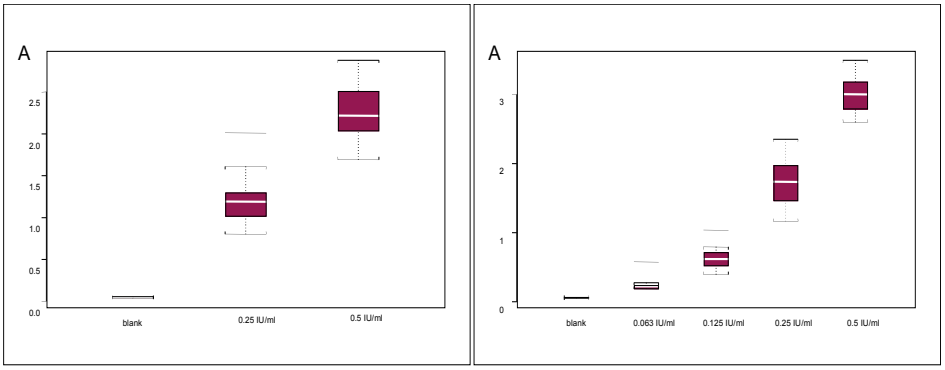


Figure. 5.2.2: Boxplots with absorbance (A) values of 20- replicates (left) or 12 replicates (right) of WHO-LPS spikes in saline at various concentrations. (readout of the IL-6 ELISA).

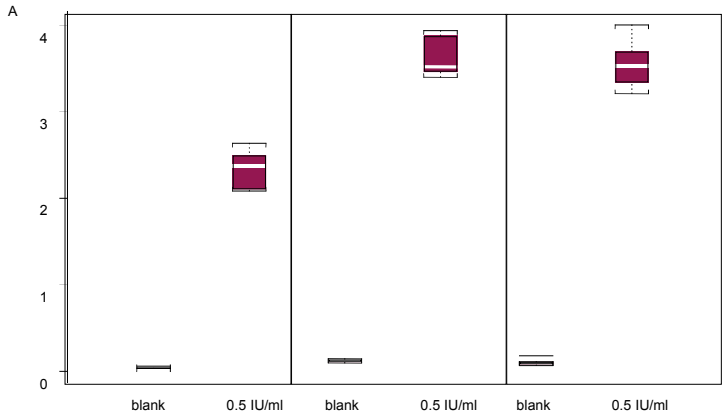


Figure. 5.2.3: Boxplots absorbance (A) values of the response of 3 different blood donations from one healthy volunteer on consecutive days with WHO-LPS (IU/ml) in saline at 0.0 IU/ml (Blank) or 0.5 IU/ml (readout of the IL-6 ELISA).

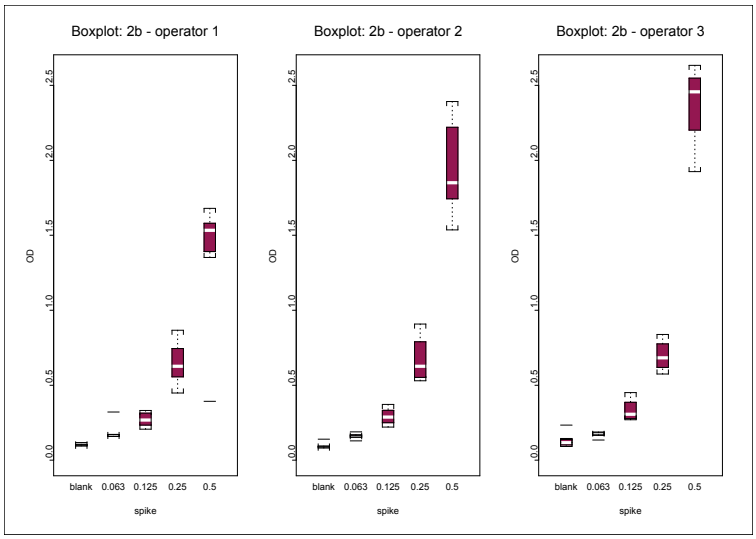


Figure. 5.2.4 A: Boxplots of absorbance (OD) values of WHO-LPS (IU/ml) in saline at 0.0 (Blank), 0.25 IU/ml (S0.25) or 0.5 IU/ml. (S0.5) with 3 different operators (readout of the IL-6 ELISA).

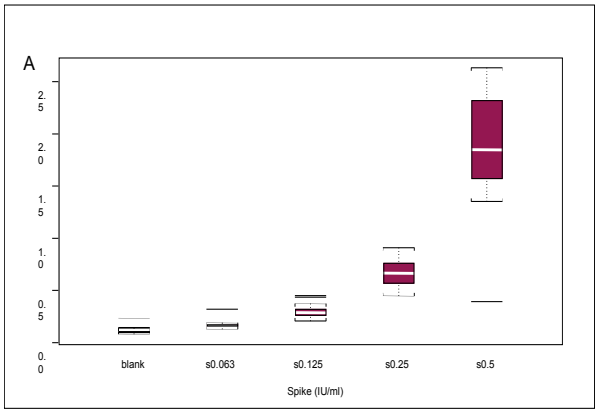


Figure. 5.2.4 B: Boxplot of absorbance (A) values of WHO-LPS (IU/ml) in saline at 0.0 (Blank), 0.25 IU/ml (S0.25) or 0.5 IU/ml. (S0.5). Combined data of Fig. 5.2.4.A (readout of the IL-6 ELISA).

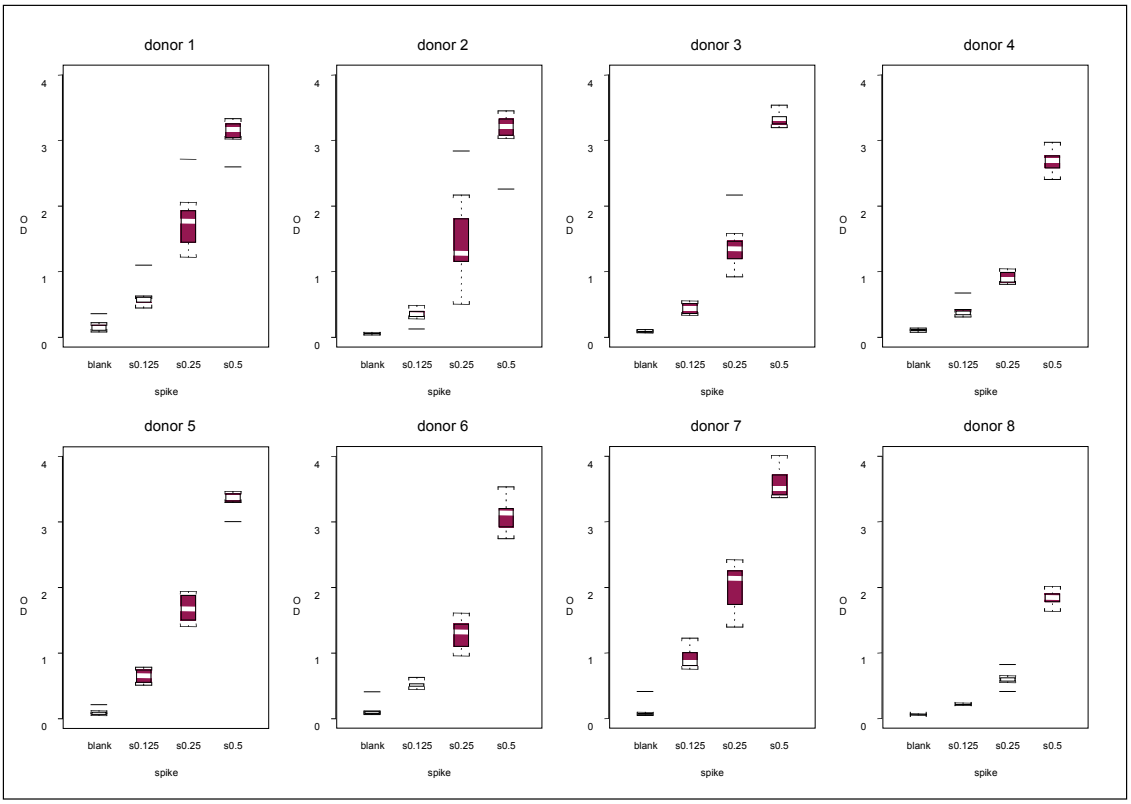


Figure. 5.2.5: Boxplots of absorbance values (OD) of the response of 8 individual donors to WHO-LPS (IU/ml) in saline at 0.0 IU/ml (blank), 0.125 IU/ml (s0.125), 0.25 IU/ml (s0.25) and 0.5 IU/ml (s0.5) (readout of the IL-6 ELISA).

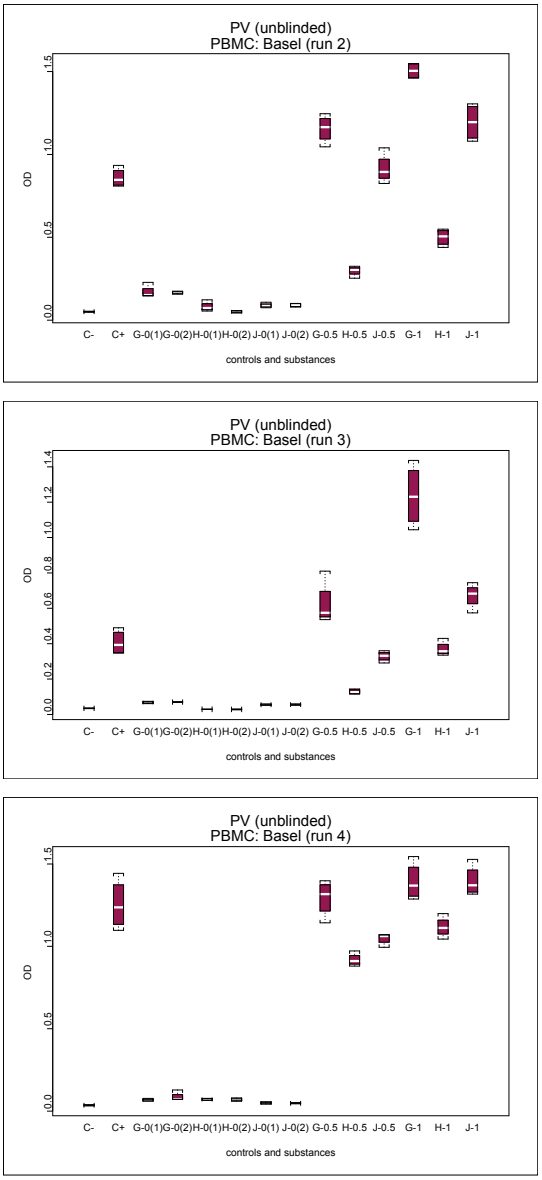


Figure. 5.2.6 A: Three different drugs were spiked with WHO-LPS at 0.0, 0.0, 0.5 and 1.0 IU/ml, respectively. Test was performed with known drugs and blinded spikes. Experiment was run 3 time independently at the Basel laboratory (readout of the IL-6 ELISA). The first run was invalidated for technical reasons (data not shown).
G = Gelafundin; J = Jonosteril; H = Heamate.
C- = negative control (saline); C+ = positive control (0.5 IU/ml in saline).

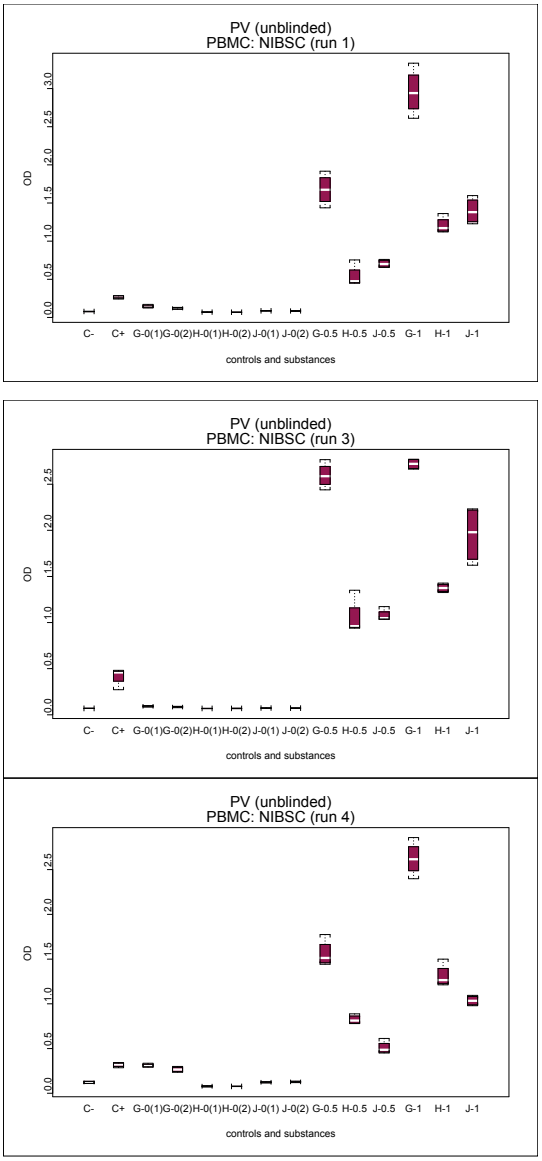


Figure. 5.2.6 B: Three different drugs were spiked with WHO-LPS at 0.0, 0.0, 0.5 and 1.0 IU/ml, respectively. The spikes were blinded. Experiment was run 3 time independently at the laboratory of the NIBSC (readout of the IL-6 ELISA).
G = Gelafundin; J = Jonosteril; H = Heamate.
C- = negative control (saline); C+ = positive control (0.5 IU/ml in saline).

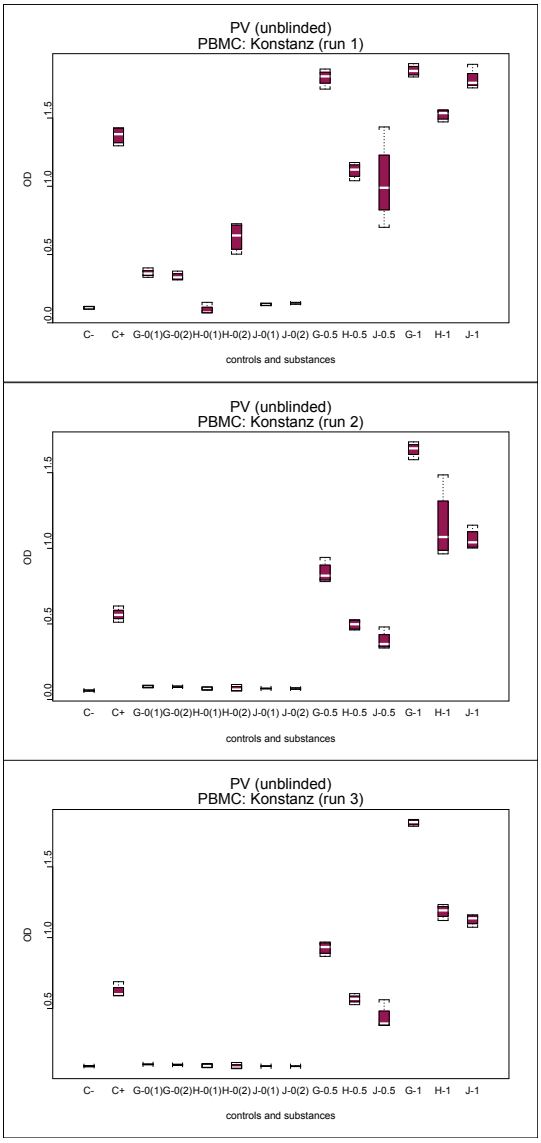


Figure. 5.2.6 C: Three different drugs were spiked with WHO-LPS at 0.0, 0.0, 0.5 and 1.0 IU/ml, respectively. The spikes were blinded. Experiment was run 3 time independently at the Konstanz laboratory (readout of the IL-6 ELISA).
G = Gelafundin; J = Jonosteril; H = Heamate.
C- = negative control (saline); C+ = positive control (0.5 IU/ml in saline).

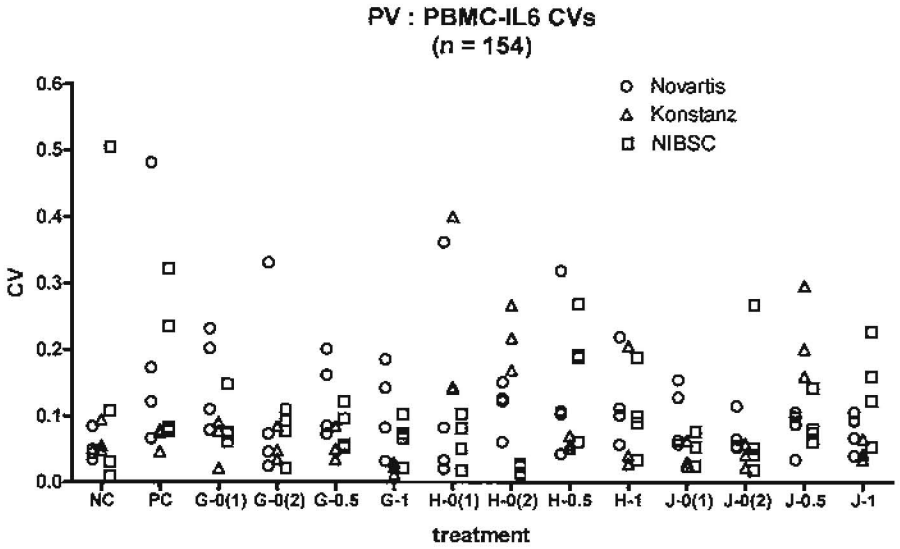


Figure. 5.2.7: Coefficient of variation (CV) of different WHO-LPS spikes (0.0, 0.0, 0.5 and 1.0 IU/ml, respectively).from the experiments as shown in fig. 5.2.6 A-C.
G = Gelafundin; J = Jonosteril; H = Heamate.
NC = negative control (saline); PC is positive control (0.5 IU/ml in saline).

5.3 Statistics

Describe the statistical approach used to evaluate the data resulting from studies conducted with the proposed test method.

A generally applicable analytical procedure was employed. This procedure includes a universal PM as well as quality criteria. First, a two-step procedure consisting of a variance-criterion and an outlier-test was applied. For this, the Dixon's test (Barnett and Lewis, 1984), which is USP approved, was chosen with the significance level of α=0.01 and applied to identify and eliminate aberrant data.

Next, the negative and the respective positive control are compared to ensure a suitable limit of detection. For this, a one-sided t-test with a significance level of α=0.01 is applied to the ln-transformed data to ensure that the response to the positive control is significantly larger than that of the respective negative control.

Finally, the samples are classified as either negative or positive by the outcome of a one-sided version of the t-test, which is based on the assigned pyrogen threshold value. The final results will be given in 2 x 2 contingency tables (table 5.3.1). These tables allow for estimation of accuracy (sensitivity and specificity) and reproducibility of the proposed test method.

Table 5.3.1: 2x2 contingency table.

		pre-defined class ("truth")		Σ
		1	***0***	
Classification by test system and PM	***1***	***a***	***b***	$a+b = n_{.1}$
	0	***c***	***d***	$c+d = n_{.0}$
Σ		$a+c = n_{1.}$	$b+d = n_{0.}$	***n***

Accuracy:

The most important statistical tool to determine accuracy (specificity and sensitivity) is the so-called PM (Hothorn, 1995). In general, it is a statistical model, which classifies a given drug by an objective diagnostic or deciding rule. The objective of a dichotomous result requires a clear cut PM, which assigns a drug in one of the two classes "pyrogenic for humans" and "non-pyrogenic for humans". Since a threshold pyrogen value will be used, a one-sided test is appropriate for the task. Because the data are normalised by a ln-transformation, a t-test is chosen. Although the variances over the range of concentration converge by the transformation, the assumptions of equal variances do generally not hold true, because it depends on additional covariates. Therefore, the one sided Welch-t-test (Snedecor and Cochran, 1989) is applied. Due to the safety aspect of the basic problem, the hypotheses of the test are

$$H_0 : \mu_{S_i j} > \mu_{S+} \quad vs \quad H_1 : \mu_{S_i j} < \mu_{S+},$$

where $\mu_{...}$ denotes the parameter of location of the respective ln-transformed distribution. This approach controls the probability of false positive outcomes directly by means of its significance level α, which is chosen as 0.01, because is assumes hazard, respectively pyrogenicity, of the tested drug in H_0, and assures safety, i.e. non-pyrogenicity. The test statistic is

$$T_{S_i j} = \frac{\bar{x}_{S+} - \bar{x}_{S_i j}}{\sqrt{\frac{s^2_{S+}}{n_{S+}} + \frac{s^2_{S_i j}}{n_{S_i j}}}}.$$

The PM is built by means of the outcome of the test. Let 0 denote safety and 1 denote hazard. The classification of S_i-j is then determined by

$$S_ij = 0, \text{ if } T_{S_ij} > t_{0.99;n_{S+}+n_{S_ij}-2},$$

$$S_ij = 1, \text{ else,}$$

where $t_{0.99;n_{S+}+n_{S_ij}-2}$ the 0.99-quantile of the t-distribution with $n_{S+} + n_{S_ij} - 2$ degrees of freedom. The number of replicates for every control and sample, i.e. $n_{...}$, was harmonised to be four. Due to the possibility of removing one observation by the outlier test, the number of replicates could be reduced to three. The classification of a version of a drug is regarded as an independent decision. Therefore, the niveau α is local.

Finally, the classifications of the drugs will be summarised in 2x2 contingency table (table 3). From these tables, estimates of the sensitivity (S_E), i.e. the probability of correctly classified positive drugs and specificity (S_P), i.e. the probability of correctly classified negative drugs, will be obtained by the respective proportions. Where

$$S_E = a / (a + c) * 100\%$$
and
$$S_P = d / (b + d) * 100\%.$$

Furthermore, these estimates will be accompanied by confidence intervals, which will be calculated by the Pearson-Clopper method [15]. For example, let $\hat{p}_{SE}$ denote the proportion, namely the sensitivity, under investigation. Then the confidence interval to a niveau α is calculated as

$$\left[p_{SE}^{L} = \frac{a F_{2a;2(n_{1.}-a+1);\frac{\alpha}{2}}}{n_{1.} - a + 1 + a F_{2a;2(n_{1.}-a+1);\frac{\alpha}{2}}}; p_{SE}^{U} = \frac{(a+1) F_{2(a+1);2(n_{1.}-a);1-\frac{\alpha}{2}}}{n_{1.} - a + (a+1) F_{2(a+1);2(n_{1.}-a);1-\frac{\alpha}{2}}} \right],$$

where $F_{...}$ denotes the respective quantile of the F-distribution and $n_{1.}$ is the sample size of the positive drugs and a the number of correctly classified drugs.

By contaminating the drugs artificially and by defining a threshold value, which is assumed to be appropriate, the class of a drug is determined beforehand. The versions of drugs, which are effectively contaminated, but below the threshold dose, are considered to be negative, respectively safe, because their contamination is not crucial for humans in terms of ELC.

Reproducibility:

The analysis of the intra- and interlaboratory reproducibility was assessed from the three identical and independent runs conducted in each of 3 laboratories. The comparison of

the three runs was carried out blindly such that the testing facility did not know the true classification of the sample, either pyrogenic or non-pyrogenic. By this procedure only the randomness and reproducibility of the methods was assessed and not systematic errors, which may have arisen from other sources, e.g. logistics or preparation of the samples. This means that if a sample was (mis)classified in all three runs the result is reproducible regardless of the (mis)classification of the sample. Therefore, a measure of similarity, i.e. complete simple matching with equal weights, was preferred to the coefficient of correlation for 2x2 contingency tables.
The study was designed as follows: each laboratory had to conduct three independent runs with the same 12 samples (3 test items with 4 blinded spikes each) and two controls, i.e. saline as a negative control (C-) and a 0.5 EU/ml LPS-spike in saline as a positive controls (C+). The samples were derived from the three substances Gelafundine, Haemate and Jonosteril. Per run, each substance was blindly spiked twice with saline, once with 0.5 EU/ml LPS and once with 1 EU/ml LPS, which resulted in a balanced design with regard to positive and negative samples, i.e. samples expected to be pyrogenic and non-pyrogenic, respectively.

The three independent runs per testing facility provide the information on which the assessment of the intralaboratory reproducibility is based. The combined results of the three runs per testing facility were used to determine interlaboratory reproducibility. The correlation of the prediction (in terms if the Bravais-Pearson coefficient of correlation) between all runs is calculated, independent of whether that classification is true or false. A BP-correlation of 1 is calculated, if two runs gave exactly the same predictions for the twelve substances. If one run gives adverse classifications for all substances than the other, the correlation is –1. As these calculations do not need information of the true status of a sample, they were carried out blinded.

5.4 Tabulated results

Provide a summary, in graphic or tabular form, of the results.
See tables 5.4.1 and 5.4.2.

Table 5.4.1: Results of testing 3 substances 3 times by 3 laboratories. Classifications after applying the PM (compare to fig. 5.2.5)

Sample	DL (Novartis, Basel)			NL 1 (Konstanz)			NL 2 (NIBSC)		
	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3
G-0 (1)	0	0	0	0	0	0	0	0	1
G-0 (2)	0	0	0	0	0	0	0	0	0
H-0 (1)	0	0	0	0	0	0	0	0	0
H-0 (2)	0	0	0	0	0	0	0	0	0
J-0 (1)	0	0	0	0	0	0	0	0	0
J-0 (2)	0	0	0	0	0	0	0	0	0
G - 0.5	1	1	1	1	1	1	1	1	1
H - 0.5	0	0	0	1	1	1	1	1	1
J - 0.5	1	1	1	0	0	0	1	1	1
G - 1	1	1	1	1	1	1	1	1	1
H - 1	1	0	1	1	1	1	1	1	1
J - 1	1	1	1	1	1	1	1	1	1

"0"denotes "non-pyrogenic"; "1" denotes "pyrogenic".

Table 5.4.2: Results of the validation study of 10 drugs, spiked with WHO-LPS at 0.0, 0.25, 0.5, 0.5 and 1.0 IU/ml, respectively and tested in 3 different laboratories. Samples and spikes were blinded. Classifications after applying the PM (compare to fig. 5.2.7).

drug (code)	spike EU/ml	"truth"	results Basel	Konstanz	NIBSC
Beloc (BE)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**0**
	0.50	**1**	**0**	**1**	**1**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
Binotal (BI)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**1**	**0**
	0.50	**1**	**1**	**1**	**0**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
Ethanol 13% (ET)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**1**	**0**
	0.50	**1**	**1**	**1**	**1**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
Fenistil (FE)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**1**	**0**
	0.50	**1**	**1**	**1**	**1**

BRD: PBMC/IL-6 March, 2006

drug (code)	spike EU/ml	"truth"	results Basel	Konstanz	NIBSC
	0.50	1	1	1	1
	1.00	1	1	1	1
Glucose 5% (GL)	0.00	0	0	0	0
	0.25	0	0	0	0
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
"Drug A" 0.9% NaCl (LO)	0.00	0	0	0	0
	0.25	0	0	0	0
	0.50	1	0	1	1
	0.50	1	0	1	1
	1.00	1	1	1	1
MCP (ME)	0.00	0	0	0	0
	0.25	0	0	0	0
	0.50	1	1	1	0
	0.50	1	1	1	1
	1.00	1	1	1	1
"Drug B" 0.9% NaCl (MO)	0.00	0	0	0	0
	0.25	0	0	0	0
	0.50	1	0	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Orasthin (OR)	0.00	0	0	0	0
	0.25	0	0	0	0
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Sostril (SO)	0.00	0	0	0	0
	0.25	0	0	0	0
	0.50	1	0	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1

"0"denotes "non-pyrogenic"; "1" denotes "pyrogenic"; NA is not assessed.

5.5 Coding of data

For each set of data, indicate whether coded chemicals were tested, whether experiments were conducted without knowledge of the chemicals being tested, and the extent to which experiments followed GLP guidelines.
Blinding of drugs and/or spikes is indicated with the data.

5.6 Circumstances

Indicate the "lot-to-lot" consistency of the test substances, the time frame of the various studies, and the laboratory in which the study or studies were conducted. A coded designation for each laboratory is acceptable.
In each part of the study, all samples are derived from one (clinical) lot.

5.7 Other data available

Indicate the availability of any data not submitted for external audit, if requested.
All relevant data were submitted with the present BRD.

6 Test Method Accuracy

6.1 Accuracy

Describe the accuracy (e.g., concordance, sensitivity, specificity, positive and negative predictivity, false positive and negative rates) of the proposed test method compared with the reference test method. Explain how discordant results in the same or multiple laboratories from the proposed test were considered when calculating accuracy.
Test method accuracy was assessed in two large scale experiments performed with the drugs outlined in table 3.3.1 and table 3.3.2 in section 3 respectively. As described before one experiment was performed in an early stage of the study with 3 different drugs, tested 3 times and the other final experiment all drugs were tested once in the three participating laboratories. From the first experiment a preliminary estimate of sensitivity and specificity can be figure out, whereas the second is regarded as the established accuracy for the PBMC/IL-6 assay.

6.1.1 Preliminary estimate of the accuracy of the PBMC/IL-6 test. the study a different concept for interference testing was used. The developing laboratories determined for each drug (outlined in table 3.3.2, section 3.3) the smallest dilution within the MVD that showed no interference or an acceptable degree of interference with the spike recovery. In general the lowest dilution of the sample allowing for a 50-200% spike recovery was chosen. In addition, the positive control (PC) set at 0.5 EU/ml saline was used as the classification threshold. The laboratory procedure as described in the protocol was maintained throughout the study. Although it was realized there were some drawbacks to the concept for interference testing and applying the PC as a threshold, this small scale study allows for a preliminary estimate of the accuracy of the PBMC/IL-6 method.
It has to be noted that this part of the study was designed to provide an estimate of the intra- and interlaboratory reproducibility. Therefore it will also be discussed in detail in section 7 (Test Method Reliability).

According to the PM applied during an early phase of the study the outcome (positive/negative) is related to the positive control (PC=0.5 EU/ml). If absorbance of sample > absorbance of PC, then the sample is classified as positive. If absorbance of sample < PC, then the sample is classified as negative. While performing the experiments during this phase it was realized that there is a flaw between interference testing and the PM. After the interference testing, the lowest dilution of the sample allowing for a 50-200% spike recovery was chosen. It is obvious that even with a flawless repeatability of the assay; a spike recovery between 50%-100% would be classified as negative according to the preliminary PM. In addition, due to unforeseen problems with the preparations of the spike, the recovery of the spikes was far below 100%. (This is outside the scope of the study and will not be discussed). As a consequence of the employed preliminary setup of the study the sensitivity will be underestimated, and the specificity will be overestimated.

In short, three test substances were spiked with WHO-LPS at four levels (0, 0, 0.5 and 1.0 EU/ml respectively), which resulted in 12 samples with a balanced design with regard to positive and negative samples (i.e. samples expected to be pyrogenic and non-pyrogenic respectively. These 12 samples were three times tested in three laboratories. In total there were 108 classifications from 12 samples in 3 runs and in 3 laboratories (3x3x12=108). Results are described in detail in section 7. A 2x2 contingency table was constructed (table 6.1.1), from which the estimates of sensitivity and specificity can easily be derived.

Table 6.1.1: 2x2 contingency table. The PM applied to a preliminary study.

		True status of samples		**Total**
		+	-	
PM	+	47	1	**48**
	-	7	53	**60**
	Total	**54**	**54**	**108**

The specifications of specificity and sensitivity described in section 5.3 were applied to these results.
The specificity (Sp) of the PBMC/IL-6 assay is 98.1% (53/(1+53)*100%), 95% confidence interval [0.901; 0.999]. The sensitivity (Se) equals 87% (47/(47+7)*100%), 95% confidence interval [0.751; 0.946]. As outlined previously the specificity is overestimated and the sensitivity is underestimated as a result of the design of this part of the study.

6.1.2 Test method accuracy of the proposed PBMC/IL-6 method. To assess accuracy of the proposed method, 10 substances (listed in table 3.1.1, section 3) were spiked with five different concentrations of the WHO-LPS (one of which is negative). Thus, in total, 50 samples have been tested in each laboratory.
To permit the application of the chosen PM, each drug was diluted to its individual MVD, which has been calculated using the ELC to that drug (listed in section 3). Lesser dilutions were tested by the DL, and showed no interference. Therefore interference was not expected at the individual MVD. Each substance had their own specific set of endotoxin spike solution, ensuring that after spiking the undiluted drug, it contained 0.0; 0.25; 0.5; 0.5 and 1.0 EU/ml respectively when tested at its MVD. A detailed description of the sample preparation was supplied to the three independent laboratories (shown in table 5.1.1 for convenience). To put more weight to the result of this part of the validation, the spikes were blinded and coded by QA ECVAM. The raw data and the graphical presentation of these raw data are shown in the section 5 (table 5.4.2).
Accuracy was assessed by applying the PM to the results (summarized in table 5.3.2) and evaluating the concordance in this section in a two by two contingency table (table 6.1.2). As described above 10 substances, spiked with 5 different WHO-LPS concentrations were tested in three laboratories and consequently a maximum of 150 data were available for analysis.

As intralaboratory reproducibility was successfully shown in previous experiments (analyzed in section 7), only one run performed in each laboratory was considered sufficient.

Table 6.1.2: 2x2 contingency table. PM applied to the PBMC/IL-6 test result of 10 different substances assessed in three different laboratories.

		True status of samples +	-	Total
PM	+	83	3	**86**
	-	7	57	**64**
	Total	**90**	**60**	**150**

All the 150 available sets of replicates met the quality criteria as described in section 5.3. The specificity and sensitivity of the PBMC/IL-6 method was estimated as described in section 5.3.

The specificity of the PBMC/IL-6 assay is 95.0% (57/(3+57)*100%), 95% confidence interval [0.861;0.990]. The sensitivity equals 92.2% (83/(83+7) *100%), 95% confidence interval [0846;0.968]. (See table 6.1.3). The specificity varied from 85% up to 100% within the three laboratories, and the sensitivity varied from 83.3% up to 100%.

Table 6.1.3: Specificity and sensitivity of the PBMC/IL-6 assay

	N total	N correctly identified	proportion	95% CI lower limit	95% CI upper limit
Specificity (Sp)	60	57	95.0%	86.1%	99.0%
Sensitivity (Se)	90	83	92.2%	84.6%	96.8%

6.2 Concordancy to *in vivo* reference method

Discuss results that are discordant with results from the in vivo reference method.
Not applicable.

6.3 Comparison with reference methods

Discuss the accuracy of the proposed test method compared to data or recognized toxicity from the species of interest (e.g., humans for human health-related toxicity testing), where such data or toxicity classification are available. This is essential when the method is measuring or predicting an endpoint for which there is no preexisting method. In instances where the proposed test method was discordant from the in vivo reference test method, describe the frequency of correct predictions of each test method compared to recognized toxicity information from the species of interest.
Not applicable.

BRD: PBMC/IL-6 March, 2006

6.4 Strength and limitations

State the strengths and limitations of the proposed test method, including those applicable to specific chemical classes or physical-chemical properties.
It appears the proposed test is applicable to most classes of medicinal products, at least those that are non- or low-toxic to cells *in vitro*. In addition, the test may be employed to assess pyrogenicity of various medical devices, such as (biological) bovine collagen bone implants.

6.5 Data interpretation

Describe the salient issues of data interpretation, including why specific parameters were selected for inclusion.
No issues.

6.6 Comparison to other methods

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the results obtained with both test methods should be compared with each other and with the in vivo reference test method and/or toxicity information from the species of interest.
Not applicable.

7 Test Method Reliability (Repeatability/Reproducibility)

7.1 Selection of substances

Discuss the selection rationale for the substances used to evaluate the reliability (intra-laboratory repeatability and intra- and inter-laboratory reproducibility) of the proposed test method as well as the extent to which the chosen set of substances represents the range of possible test outcomes.

The rationale for the selection of the substances is described in section 3.3. In short: for the present studies endotoxin (WHO-LPS) was selected as the model pyrogen, since it is well defined biological standard and readily available. Selected test substances were medicinal products available on the market. These batches are released by the manufacturers and comply with the Marketing Authorisation file and European Pharmacopoeia. Therefore these batches are considered to contain no *detectable* pyrogens. To test the method reliability the medical products were spiked with endotoxin.

7.2 Results

Provide analyses and conclusions reached regarding the repeatability and reproducibility of the proposed test method. Acceptable methods of analyses might include those described in ASTM E691-92 (13) or by coefficient of variation analysis.

In an early phase of the study, the intralaboratory repeatability and reproducibility of the test method was assessed in a series of experiments conducted in the DL). Series of blanks (saline, 0 EU/ml) and spikes of WHO-LPS in saline were tested. These experiments (1A, 1B, 2A, 2B and 2C) are summarized in table 7.2.1.

Table 7.2.1: Summary of experiments with WHO-LPS in saline.

Experiment	Spikes (EU/ml) in saline	n (per spike)	Repetitions of experiment	N
1A	0; 0.25; 0.5	20	1	60
1B	0; 0.063; 0.125; 0.25; 0.5	12	1	60
2A	0; 0.5	8	3	48
2B	0; 0.063; 0.125; 0.25; 0.5	8	3	120
2C	0; 0.125; 0.25; 0.5	8	8	256

The data were used to answer questions regarding the nature of the distribution, the variance and its behavior over the range of response in replicated measurements under identical conditions. In addition, reliability of the test method was assessed by the maintenance of minimum criteria, e.g. a detection limit below an a-priori chosen positive control or a dose-dependent standard curve (table 7.2.1, experiment 1B). With the data of this experiment an assessment of the limit of detection of the test can be done by calculating the smallest spike, which can be discriminated from the blank.

The second group of experiments was meant to analyze the variation in detail. For this purpose the major sources variation were assessed separately, i.e. behavior of a donor in time (experiment 2A), operator (exp. 2B) and different donors (exp.2C). A total of 554 observations were collected and analyzed.

First the shape of the distribution at a spike was assessed (not shown). Most of the data showed normal-distribution.
Based on the experience that there is a monotone increasing relationship between the mean-responses and the variation (empirical variance or standard deviation), the analysis focuses on the CV. Outliers were removed with the Grubbs-test (α =5%). The CV should be distributed symmetric around a constant factor, if the mean-variance relationship is linear. A plot of all CVs against their corresponding means is shown in figure 5.2.1. From the figure it is clear that at this stage of the study, the CV for most sets of replicates is lower than CV 0.25, only seven out of the 61 sets showed a CV above 0.25. This higher variation originates from only one of the experiments (experiment 2C), with 5 of the 7 higher values observed for donor 1 and donor 2.

The outliers were identified on the assumption of normally distributed data as well as a log-normal distribution. At this point the Grubbs-test was chosen and the kind of outlier (lower or upper) and the significance level α (5% and 1% significance level) were recorded. Altogether there were 19 outliers identified out of 544 observations, which is a proportion of only 3.5%. Most of the outliers originate from blanks, six out of the seven were upper outliers. This might be explained by small inhomogeneities, which will have a major impact. In addition, the raw data (plate-readouts) showed no obvious edge-effects or trends.

The results of test 1A (figure 5.2.2) show a low variation between the 20 replicates of each spike. The figure shows one outlier, which could detected with the Grubbs-test. The spike of 0.25 EU/ml was easily detected. Test 1B showed a dose response curve with 12 replicates for each spike concentration. Two outliers could be identified. It is obvious that that even a 0.0625 EU/ml of LPS standard can be detected with the PBMC/IL6 assay.

Test 2A (figure 5.2.3) was included to assess the behavior of a donor in time. Blood samples were collected at day 1, day 36 and day 41 respectively. Data are presented in figure 5.2.3. Unfortunately, the interpretation of the result is impaired because the first experiment was executed by another operator than the last two experiments. It is not clear whether the lower spike -response of the first day is due to the operator or to the day. However, in general the results of the three days are comparable and the 0.5 EU/ml spike can be detected beyond doubt.

Five LPS spikes were tested by three operators in parallel while conducting experiment 2B. The results, presented in boxplots in figure 5.2.4, are similar. Only the level of the readout of the 0.5 EU/ml spike differ. Again the 0.063 EU/ml spike was detected by every operator. It can be concluded that the operator does not have an important impact on the performance of the assay.

The last experiment 2C was designed to show the robustness of the assay with respect to different donors. Therefore 8 donors were involved and for each donor eight replicates of each of the spikes (0; 0.125; 0.25; 0.5 EU/ml) were generated. Figure 5.2.5. shows that the level of the OD-values differ from donor to donor due to their individual sensitivity to

LPS. But every donor reacts to the lowest spike tested, the 0.125 EU/ml-spike. As mentioned before a few very large variances, are noticeable for donor 1 and 2. Regarding the overall low variances, this is considered an incident. The donor will have an impact on the results of an assay in terms of reproducibility over donors, because the data differ stronly between donors. But in terms of a qualitative decision all donor behaved the same. To minimise the influence of different donors, acceptance criteria have already been included in the protocol for PBMC/IL-6 assay.

In conclusion:
The data showed a very stable statistical properties over all the experiments. The shape of replicates, the numbers of outliers and the variation reveal a general structure. Nevertheless the variation can be quite high. Even if the outliers are removed, 25% of the CVs are larger than 0.2 and 10% are over 0.25. This is an acceptable amount of variation, but this might be higher if the number of replicates is reduced to four.
The sensitivity of the assay is very high. Even the spike of 0.063 EU/ml can be detected with low errors-rates (false-positive, false-negative). The two covariates "day" and "operator" have just a small impact on the result of the assay. The influence of the covariate "donor" cannot be neglected with regard to reproducibility of the OD. Nevertheless, for every donor the spikes could be discriminated from each other. Therefore the intralaboratory repeatability is considered satisfactory. The 3-4% percentage outliers, as determined by the Grubbs test is considered acceptable.

Intra- and interlaboratory reproducibility.
After transfer of the PBMC/IL-6 assays to two other laboratories, a dose response experiments was performed by all three laboratories. For this study 7 concentrations were tested in a dose response curve (0, 0.125, 0.25, 0.5, 1, 2, 4 EU/ml, at least 8 replicates). A participating laboratory qualified for taking part in the next part of the study by producing a dose response curve, with a limit of detection of at least 0.5 EU/ml and a CV < 0.4 (data not shown).

Testing 3 different medicinal substances, Gelafundin, Jonosteril and Haemate (described in table 3.3.2, section 3.3.) assessed the intra- and interlaboratory reproducibility. Test substances and their spikes were appropriately blinded. Test substances were tested, at a predefined dilution above the MVD, independently in 3 different laboratories, 3 times each. The three test substances were spiked with WHO-LPS at four levels (0, 0, 0.5 and 1.0 EU/ml respectively), which resulted in 12 samples with a balanced design with regard to positive and negative samples (i.e. samples expected to be pyrogenic and non-pyrogenic respectively. In addition a negative control (saline) and positive control (0.5 EU/ml in saline) were included to establish assay validity. To avoid interference, the DL performed interference testing in terms of the BET, i.e. 50-200% spike recovery, and decided on the dilution of the test substances. Dilutions chosen for Gelafundine, Haemate, Jonosteril were 1:4, 1:2 and 1:1 respectively. The results are graphically presented using the absorbance values of the three runs (shown in section 5, fig. 5.2.5).

BRD: PBMC/IL-6 March, 2006

From the experiment with LPS-WHO only it was concluded that CV for the PBMC/IL-6 assay is < 0.4, which is acceptable. It was envisaged that the CV was likely to be higher when testing different substances (different matrices) and was assessed for the current set of data. A plot of all CVs for all sets of 4 replicates of a spiked drug is shown in figure 5.2.7. From the figure it is clear that the CV for a set of 4 replicates of one spike concentration is usually below 0.45, which is considered acceptable for a biological assay. Only one set of data showed an exceptional high (CV>1.1) which is probably due to a pipetting error. For the remainder of the studies the CV criteria applied as validity criteria of the PBMC/IL-6 assays was arbitrarily set at CV<0.45.

The analysis of the intralaboratory reproducibility was assessed from the three identical and independent runs conducted in each laboratory. The comparison of the three runs was carried out blindly such that the laboratory did not know the true classification of the sample (either pyrogenic or non-pyrogenic). By this procedure only the randomness and reproducibility of the methods was assessed and not systematic errors, which may have arisen from other sources, e.g. logistics or preparation of the samples. This means that if a sample was misclassified in all three runs the result is 100% intralaboratory reproducible (regardless of the misclassification of the sample).

According to the preliminary PM applied during this phase of the study the outcome (positive/negative) was related to the positive control (PC=0.5 EU/ml). If absorbance of sample > absorbance of PC, then sample is classified as being positive. If absorbance of sample < PC, sample is classified as negative (positive/pyrogenic = 1, negative/non-pyrogenic = 0).
During this phase it was realized that there is a flaw between interference testing and the PM. After the interference testing, the lowest dilution of the sample allowing 50-200% spike recovery was chosen. It is obvious that even with a flawless repeatability of the assay, a spike recovery between 50%-100% would be classified as negative according to the preliminary PM.

From the three independent runs summarized in table 5.4.1, the intralaboratory reproducibility can be calculated for the separate laboratories (table 7.2.2). For these calculations there is no need for information of the true status of the sample. A minimum criterion for the establishment of an assay is that experiments carried out with the same samples should result in a high concordance of classifications.

Table 7.2.2: Intralaboratory reproducibility, assessed by correlation between different runs. Result of testing 3 substances 3 times by 3 laboratories.

	DL (Basel)	**NL1 (Konstanz)**	**NL2 (NIBSC)**
Run 1 - Run 2	92% (11/12)	100% (12/12)	100% (12/12)
Run 1 - Run 3	100% (12/12)	100% (12/12)	92% (11/12)
Run 2 - Run 3	92% (9/12)	100% (12/12)	92% (11/12)
Mean	94%	100%	94%
Proportion showing the same result in 3 runs	92% (11/12)	100% (12/12)	94% (11/12)

BRD: PBMC/IL-6 March, 2006

Each of the assays performed by the laboratories fulfilled the sensitivity criterion, i.e. the assays showed a significant difference between C- and C+. All results could be included in the analysis. From table 7.2.2 it can be read that the between runs reproducibility ranges from 92 to 100%. The mean intralaboratory reproducibility is very good (94 to 100%) for all three participating laboratories.

The interlaboratory reproducibility of the PBMC/IL-6 method was assessed in a similar manner to the intralaboratory reproducibility. A summarizing method to combine the three runs per laboratory is considered not appropriate, because it would mask misclassification. Therefore each run of one laboratory was compared with all runs of another laboratory. This results optimally in 108 comparisons between the data sets of two laboratories. The measure of similarity is then the proportion of equally classified samples. These proportions are summarized in table 7.2.3, show that there is a good interlaboratory reproducibility varying from 81% to 89% (overall mean: 85%).

Table 7.2.3: Interlaboratory reproducibility: assessed by inter-laboratory correlations. Result of testing 3 substances 3 times by 3 laboratories.

Laboratories	Interlaboratory Reproducibility	Number of equal predictions
DL – NL1	81%	87 / 108
DL – NL2	86%	93 / 108
NL1 – NL 2	89%	96 / 108
Mean	85%	

DL = Basel; NL1 = Konstanz; NL2 = NIBSC

Also from the result of the large-scale study (testing 10 substances spiked with 5 separate spikes), the interlaboratory reproducibility can be estimated (table 7.2.4). The reproducibility varied from 84% to 90% between two laboratories. All three laboratories found the same result for 40 out of 50 samples (equals 80%).

Table 7.2.4: Interlaboratory reproducibility: Assessed by testing of 10 substances, spiked 5 times. One run of 50 samples by three different laboratories.

Laboratories	Interlaboratory Reproducibility	Number of equal predictions
DL - NL1	84%	42 / 50
DL - NL2	86%	43 / 50
NL1 – NL2	90%	45 / 50
Mean	87%	
same result in all labs	80%	40 / 50

DL =Basel; NL1 = Konstanz; NL2 = NIBSC

Conclusion: It is shown that the mean intralaboratory reproducibility, assessed by the proportion of equally classified samples between different runs varies from 94% to 100%

between the three participating laboratories. The interlaboratory reproducibility between two laboratories varied from 81% to 89% in one large scale blinded experiment and from 84% to 90% in the other large scale blinded experiment. All three participating laboratories predicted the same in 80% of the measurements. It has to be noted that a substantial part of the samples was 0.5 EU/ml and therefore close to the defined pyrogenicity-threshold of the PBMC/IL-6 assay

7.3 Historical data

Summarize historical positive and negative control data, including number of experiments, measures of central tendency, and variability.
Not applicable.

7.4 Comparison to other methods

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the reliability of the two test methods should be compared and any differences discussed.
Not applicable.

8 Test Method Data Quality

8.1 Conformity

State the extent of adherence to national and international GLP guidelines (7-12) for all submitted data, including that for the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method. Information regarding the use of coded chemicals and coded testing should be included.
The studies were done in concordance with the guidelines for GLP. Written protocols and approved standard operating procedures were followed during the entire course of the study. Deviations were recorded and, where appropriate, approved in amendments. All data are stored and archived. As mentioned, samples were appropriately blinded.

8.2 Audits

Summarize the results of any data quality audits, if conducted.
No audits were conducted.

8.3 Deviations

Discuss the impact of deviations from GLP guidelines or any noncompliance detected in the data quality audits.
Not applicable.

8.4 Raw data

Address the availability of laboratory notebooks or other records for an independent audit. Unpublished data should be supported by laboratory notebooks.
All records are stored and archived by the contributing laboratories and available for inspection.

9 Other Scientific Reports and Reviews

9.1 Summary

Summarize all available and relevant data from other published or unpublished studies conducted using the proposed test method.

Relevant data obtained with the proposed method are described in a number of published studies and reports, which are added in the Appendix B.

An *in vitro* monocyte activation test that detected pro-inflammatory and pyrogenic contaminants, was first applied some 15 years ago (Poole et al., 1988). A number of variants of the original test system have since been described, although the underlying principle of each variant remains the same. The test preparation is cultured with monocytes, either as peripheral blood mononuclear cells, PBMCs, diluted whole blood or cells of a monocytoid cell line such as MONO MAC-6 (MM6). Contaminants in the test sample activate CD14/TLR receptors, which stimulate the release of an endogenous pyrogenic cytokine from the monocytes (Poole and Gaines Das, 2001).

Early studies report on opimization of the test method, e.g. improving the lower limit of detection, incubation times and cytokine readout, using model pyrogens such as LPS. Limited information is available on the actual testing of medicinal products.

Most interestingly, Taktak et al (1991) described several batches of a medicinal product (serum albumin) that caused adverse (pyrogenic) reactions in recipients. These lots were not detected by either BET or rabbit test but only by the *in vitro* monocytoid cell test.

In a study using whole blood and monocytoid cell lines as the sources of monocytoid cells (Nakagawa et al., 2002) it was reported that the structurally diverse pyrogens endotoxin, peptidoglycan, Staphylococcus aureus, Cowan 1, Curdlan and poly(I.C) all stimulated the release of cytokines.

The cytokine readout included tumour necrosis factor (TNF), interleukin-1 (IL-1), and IL-6 (reviewed by Poole and Gaines Das, 2001 and Poole et al., 2003). Other cytokines, e.g. IL-8, are also produced in large quantities in response to pyrogenic contaminants but their roles in fever are less well studied. The preferred readout is usually IL-6 because IL-6, unlike IL-1 and TNF, is secreted entirely into the cell-conditioned medium in large quantities, thereby permitting its complete estimation (Poole et al, 1988; Poole et al., 1989, Taktak et al., 1991).

Also, certain pro-inflammatory bacterial components stimulate the production of IL-6 but not TNF and IL-1 (Reddi et al., 1996), and IL-6 induction via Toll-like (pyrogen) receptors rapidly follows the recognition of microbial products (Pasare and Medzhitov, 2003).

It is stressed throughout these studies using whole blood and PBMCs that only healthy donors not taking any medication must be used for testing pharmaceuticals. Various drugs such as cortisone suppress interleukin release and would therefore exclude a blood donor from the experiment. An infection can stimulate release, as reflected in an increased baseline value or an abnormally high interleukin response. Therefore, the PBMC/IL-6 test may only be used if samples have first been shown not to cause

interference. There is no indication that the blood group of the human donors influences the results of the assay.

9.2 Discussion

Comment on and compare the conclusions published in independent peer-reviewed reports or other independent scientific reviews of the proposed test method. The conclusions of such scientific reports and reviews should be compared to the conclusions reached in this submission. Any ongoing evaluations of the proposed test method should be described.

The validation described in this BRD is the first time that such an extensive study for specificity and accuracy using actual medicinal products spiked with endotoxine is carried out. Hence, there are no comparing reports in independent peer-reviewed journals available.

9.3 Results of similar validated method

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the results of studies conducted with the validated test method subsequent to the ICCVAM evaluation should be included and any impact on the reliability and accuracy of the proposed test method should be discussed.

As mentioned, *in vitro* monocyte activation test methods for the detection of pyrogenic contaminants are being developed over the course of the past two decades. A number of variants have been described, although the underlying principle of each variant remains the same. The test preparation is cultured with monocytoid cells, either as peripheral blood mononuclear cells, PBMCs, (diluted) whole blood or cells of a monocytoid cell line such as MM6. Accuracy and specificity of these test methods are comparable, but in general methods using whole blood, PBMCs and the MM6 cell line appear to perform best (Hoffmann et al, 2005b). Table 9.3.1 summarises the performance of *in vitro* methods presented in the five BRDs and Table 9.3.2 compares the *in vivo* and *in vitro* pyrogen tests regarding their strengths, weaknesses, costs, time, limitations.

However, most studies (as this one) are done with model pyrogens and as yet little experience is available in the field, e.g. as part of the final batch release test-package. Experience and thus confidence in these methods will grow once regulatory authorities approve these methods and more manufacturers start to employ them. Then, on a case by case situation, it should be determined which method is best suited for the actual situation and demonstrates to pick out the appropriate, i.e. pyrogenic batches of the medicinal product.

BRD: PBMC/IL-6 March, 2006

Table 9.3.1: Summary of the performance of in vitro pyrogen tests based on monocytoid cells (see Tables 7.2.2; 7.2.4; 6.1.3)

Test	System	Read-out	Intralaboratory reproducibility (%)	Interlaboratory reproducibility (%)	Sensitivity (%)	Specificity (%)
WB/IL-6	whole blood	IL-6	DL: 83.3 NL1: 94.4 NL2: 100	DL-NL1: 85.4 DL-NL2: 85.4 NL1-NL2: 92.0	88.9	96.6
WB/IL-1	whole blood	IL-1β	DL: 88. 9 NL1: 95.8 NL2: 94.4	DL-NL1: 72.9 DL-NL2: 81.6 NL1-NL2: 70.2	72.7	93.2
96-wells WB/IL-1 [1]	whole blood	IL-1β	-	DL-NL1: 88.1 DL-NL2: 89.7 NL1-NL2: 91.5	98.8	83.6
CRYO WB/Il-1	cryo whole blood	IL-1β	-	DL-NL1: 91.7 DL-NL2: 91.7 NL1-NL2: 91.7	97.4	81.4
KN CRYO WB/Il-1 [2]	cryo whole blood	IL-1β	-	DL-NL1: 83.3 DL-NL2: 100 NL1-NL2: 83.3	88.9	94.4
PBMC/IL6	PBMC	IL-6	DL: 94.4 NL1: 100 NL2: 94.4	DL-NL1: 84.0 DL-NL2: 86.0 NL1-NL2: 90.0	92.2	95.0
PBMC-CRYO/IL-6 [3]	PBMC	IL-6	-	DL-NL1: 96 DL-NL2: 76 NL1-NL2: 80	93.3	76.7
MM6/IL-6	MM6	IL-6	DL: 100 NL1: 94.4 NL2: 94.4	DL-NL1: 90.0 DL-NL2: 89.6 NL1-NL2: 83.3	95.5	89.8

DL = developing laboratory; NL1, NL2 = naive laboratory 1 and 2
1 = data provided in Section 13 of WB/IL-1 BRD
2 = data provided in Section 13 of CRYO WB/IL-1 BRD
3 = data provided in Section 13 of PBMC/IL-6 BRD

Table amended from Hoffmann et al 2005b; results with THP cells not included

BRD: PBMC/IL-6 March, 2006

Table 9.3.2: Comparison of the in vivo and in vitro pyrogen tests regarding their strengths, weaknesses, costs, time, limitations

	Rabbit pyrogen test	**BET / LAL**	**In vitro pyrogen test**
Test materials	Liquids	Clear liquids	Liquids, potentially cell preparations, solid materials
Pyrogens covered	All (possible species differences to humans for non-endotoxin pyrogens)	Endotoxin from Gram-negative bacteria	(probably) all
Limit of detection (LPS)	0,5 EU	0,1 EU (some variants down to 0,01 EU)	0,5 EU (validated PM), some variants down to 0,001 EU
Ethical concerns	Animal experiment	About 10% lethality to bled horseshoe crabs	Some assays: blood donation
Costs*	High (200-600$/sample)	Low (50-150$/sample)	Medium (100-350$/sample)
Time required	27 h	45 min	24-30h**
Materials not testable	Short-lived radiochemicals, anesthetics, sedatives, analgetics, chemotherapeutics, immunomodulators, cytokines, corticosteroids	Most biologicals, glucan-containing preparations (herbal medicinal products, cellulose-filtered products), lipids, microsomes, cellular therapeutics	Not known (some of the materials not testable in rabbits require adaptations)
Others	No positive or negative control included, strain differences, stress affects body temperature	Potency of LPS from different bacterial species in mammals not reflected, false-positive for glucans	Possible donor differences, need to exclude hepatitis/HIV and acute infections / allergies of donors, dedifferentiation of cell lines

* = We consulted the laboratories participating in the validation study and a consultant regarding the costs of the tests. The figures we received vary significantly depending on the facility (e.g. industry, contract laboratory, control authority), frequency of testing, specific test requirements, country, etc.

** = interference testing might increase duration by 24 hours

10 Animal Welfare Considerations (Refinement, Reduction, and Replacement)

10.1 Diminish animal use

Describe how the proposed test method will refine (reduce or eliminate pain or distress), reduce, or replace animal use compared to the reference test method.

Depending on the medicinal product, one of two animal-based pyrogen tests is currently prescribed by the respective Pharmacopoeias, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET). The rabbit pyrogen test detects various pyrogens but alone the fact that large numbers of animals are required to identify a few batches of pyrogen-containing samples argues against its use. In the past two decades, the declared intention to refine, reduce and replace animal testing, has lowered rabbit pyrogen testing by 80% by allowing to use the BET as an alternative pyrogen test for certain medicinal products. Bacterial endotoxin is the pyrogen of major concern to the pharmaceutical industry due to its ubiquitous sources, its stability and its high pyrogenicity. With the BET, endotoxin is detected by its capacity to coagulate the amoebocyte lysate from the haemolymph of the American horseshoe crab, *Limulus polyphemus*, a principle recognised some 40 years ago (Levin and Bang, 1964). In the US, *Limulus* crabs are generally released into nature after drawing about 20% of their blood and therefore most of these animals survive. However, the procedure still causes mortality of about 30,000 horseshoe crabs per year, which adds to more efficient threats of the horseshoe crab population like its use as bait for fisheries, habitat loss and pollution.

The proposed test method is an alternative for the rabbit test and the BET. By replacing the rabbit test or the BET, the lives of rabbits and horseshoe crabs are spared.

10.2 Continuation of animal use

If the proposed test method requires the use of animals, the following items should be addressed:

10.2.1 *Describe the rationale for the need to use animals and describe why the information provided by the proposed test method requires the use of animals (i.e., cannot be obtained using non-animal methods).*

Not applicable.

10.2.2 *Include a description of the sources used to determine the availability of alternative test methods that might further refine, reduce, or replace animal use for this testing. This should, at a minimum, include the databases searched, the search strategy used, the search date(s), a discussion of the results of the search, and the rationale for not incorporating available alternative methods.*

Not applicable.

10.2.3 *Describe the basis for determining that the number of animals used is appropriate.*

Not applicable.

BRD: PBMC/IL-6 March, 2006

10.2.4 *If the proposed test method involves potential animal pain and distress, discuss the methods and approaches that have been incorporated to minimize and, whenever possible, eliminate the occurrence of such pain and distress.*
Not applicable.

11 Practical Considerations

11.1 Transferability

Discuss the following aspects of proposed test method transferability. Include an explanation of how this compares to the transferability of the in vivo reference test method and, if applicable, to a comparable validated test method with established performance standards.

In general, the proposed test method is not unlike other bioassays and immunoassays that are performed routinely in many laboratories.

11.1.1 *Discuss the facilities and major fixed equipment needed to conduct a study using the proposed test method*.

No extraordinary facilities are required. General laboratory equipment and analytical instruments for performing immunoassays, e.g. microtiter plate reader and –washer, are sufficient to perform the proposed test method.

11.1.2 *Discuss the general availability of other necessary equipment and supplies*.

All supplies and reagents are readily available on the market. In contrast, availability of sufficient rabbits of adequate weight and in good health for the *in vivo* reference test is sometimes reported a limitation.

It is stressed throughout these studies using PBMC, that only healthy donors not taking any medication must be used for testing pharmaceuticals. Various drugs such as cortisone suppress interleukin release and would therefore exclude a blood donor from the experiment. An infection can stimulate release, as reflected in an increased baseline value or an abnormally high interleukin response.

11.2 Training

Discuss the following aspects of proposed test method training. Include an explanation of how this compares to the level of training required to conduct the in vivo reference test method and, if applicable, a comparable validated test method with established performance standards.

11.2.1 *Discuss the required level of training and expertise needed for personnel to conduct the proposed test method.*

The proposed test method requires personnel trained for general laboratory activities in cell biology and immunochemistry or biochemistry. Techniques they should master are not unlike cell culture (aseptic operations) and immunological techniques (especially ELISA). Such expertise is available in most if not all QC-laboratories.

11.2.2 *Indicate any training requirements needed for personnel to demonstrate proficiency and describe any laboratory proficiency criteria that should be met.*

Personnel should demonstrate that they master the execution of the test. The candidate should demonstrate to meet all the appropriate assay acceptance criteria and yield accurate results (outcome) using selected test items.

11.3 Cost Considerations

Discuss the cost involved in conducting a study with the proposed test method. Discuss how this compares to the cost of the in vivo reference test method and, if applicable, with that of a comparable validated test method with established performance standards.
Two factors contribute to the cost of the proposed test method: cost of the reagents and especially personnel.
Since the proposed test method is relatively more labour-intensive, it is estimated that the cost of the proposed test method is more then the BET or the *in vivo* reference test using rabbits. Obviously, a higher throughput of tests (runs/year) such as in a QC-laboratory of a multi-product facility or in a Contract Research Organization will significantly reduce the costs per assay.
However, especially with pharmaceuticals of biological origin, the proposed test method may be cost-effective, since these products all to often are incompatible with the BET and by their nature preclude the re-use of the rabbits.

11.4 Time Considerations

Indicate the amount of time needed to conduct a study using the proposed test method and discuss how this compares with the in vivo reference test method and, if applicable, with that of a comparable validated test method with established performance standards.
Essentially the test stretches two working days. On day one the testing materials are prepared and incubated overnight with the monocytoid cells. On the second day the amount of excreted cytokines is determined by immunoassay. The total time from start to result is approximately 24 hours.
It is thus concluded that the proposed test method will take more time when compared to the reference tests, either the rabbit test or the BET. It should be noted that rabbits are tested prior to their first use by a sham test.

12 References

List all publications referenced in the submission.

References in bold are included as hardcopies in Appendix B

Barnett, V., Lewis, T. (1984). Outliers in statistical data. 2nd edition, Chichester, US: John Wiley & Sons, pp 171-172.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Clopper, C. J., Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26: 404-413

Co Tui & Schrift, M.H. (1942). A tentative test for pyrogen in infusion fluids. Proc. Soc. Exp. Biol. Med. 49: 320-330.

Cooper JF, Levin J, Wagner Jr. HN (1971). Quantitative comparison of in vitro and *in vivo* methods for the detection of endotoxin. J Lab Clin Med 78: 138-145.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005a). Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Pertbauer A, Schindler S, et al. (2005b) Validation of novel pyrogen tests based on human monocytoid cell lines. J Immunol Methods, 298(1-2):161-173.

Hothorn, L. A. (1995). Biostatistical analysis of the 'control vs. k treatments' design including a positive control group. In: Testing principls in clinical and preclinical trials, pp 19-26, Stuttgart, Germany: Gustav Fischer Verlag.

Levin J, Bang FB (1964). A description of cellular coagulation in the Limulus. Bull Johns Hopkins Hosp 155: 337-345.

Nakagawa, Y., Maeda, H. and Murai, T. (2002). Evaluation of the *in vitro* pyrogen test system based on proinflammatory cytokine release from human monocytes: comparison with a human whole blood culture test system and with the rabbit pyrogen test. Clin. Diagn. Lab. Immunol. 9: 588-592

Pasare, C. and Medzhitov, R. (2003). Toll pathway-dependent blockade of CD4+CD25+ T cell-mediated suppression by dendritic cells. Science 299: 1033-1036.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, november 1989: 17-18.

Poole, S. and Gaines Das, R.E. (2001). Towards a 'human pyrogen test'. J. Parenter. Sci. 6: 63-66.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' *in vitro* test for pyrogens. J Immunol Methods 274: 209-220.

Reddi, K., Nair, S.P. et al (1996). Surface-associated material from the bacterium Actinobacillus actinomycetemcomitans contains a peptide which, in contrast to lipopolysaccharide, directly stimulates fibroblast interleukin-6 gene transcription. Eur. J. Biochem. 236: 871-876.

Snedecor, G. W., Cochran, W. G. (1989). Statistical methods, 8th edition, Ames, US: The Iowa State University Press.

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

13 CATCH -UP VALIDATION: Human PBMC/IL-6 in vitro Pyrogen Test with Cryopreserved Monocytoid Cells.

13.1 Rationale

The use of cryopreserved cells was investigated separately during a catch-up validation that closely followed the original trial plan (both trial plans are given in Appendix A; see also BRD CRYO WB/IL-1). Cryopreserved PBMC would allow the controlling facility to store (surplus) monocytoid cells from (individual) donors for later use, thereby reducing their waste and increasing flexibility. In addition, it allows for the collection of a larger pool of cells, increasing homogeneity of the reagent. But most of all, it would allow the blood of individual donors to be screened for the absence of adventitious agents before it is actually employed on the laboratory, boosting the safety of the test method.

13.2 Test Method Protocol Components

The method follows the original method protocol (see Appendix A), with the obvious exemption of the cryopreservation of the freshly isolated PBMCs in the presence of 10% (v/v) DMSO. In addition, a 3-donor approach was used; as compared to the original method that required PMBCs from individual 4-donors. A sample is classified as pyrogenic when at least 2 donors graded the sample as pyrogenic. Details are given in the appropriate PBMC-CRYO/IL-6 protocol (Appendix A, Detailed Protocol PBMC-CRYO/IL-6 *In vitro Pyrogen Test Using Freshly Taken or Cryopreserved PBMC (SP+PB var. Novartis; electronic file name: SOP PBMC CRYO IL-6*).

13.3 Substances Used for Validation

The same 10 parenteral drugs used to determine sensitivity and specificity (see table 3.3.1.) were used for the catch-up validation. Again, each test item was tested after spiking at its individual MVD, thus came with their own specific set of 5 endotoxin spike solutions: 0.0, 0.25, 0.5, 0.5 and 1.0 EU/ml. The test items were assessed with 5 different endotoxin levels at 3 independent test facilities, yielding a total of 150 data points, biometrically considered to be sufficient for further analysis. As all the test results qualified according to quality criteria, the maximum number was analysed.

13.4 Test Method Accuracy

To assess **accuracy** of the proposed test method 10 substances (listed in table 3.3.1), were spiked with five different concentrations of the WHO-LPS (one of which is negative). To permit the application of the chosen PM, each drug was diluted to its individual MVD, which has been calculated using the ELC to that drug (listed in table 3.3.1.) Each substance had their own specific set of endotoxin spike solution, ensuring that after spiking the undiluted drug, it contained 0.0; 0.25; 0.5; 0.5 and 1.0 EU/ml respectively when tested at its MVD. A detailed description of the sample preparation was supplied to the three independent laboratories (as shown in table 5.1.2). To put more weight to the result of this part of the validation, the spikes were blinded and coded by QA ECVAM. Accuracy was assessed by applying the PM to the results and evaluating the concordance in a two by two table.

Table 13.4.1: Results of the validation study of 10 drugs, spiked with WHO-LPS at 0.0, 0.25, 0.5, 0.5 and 1.0 IU/ml, respectively and tested in 3 different laboratories. Samples and spikes were blinded. Classifications after applying the PM (compare to fig. 5.2.7).

drug (code)	**spike EU/ml**	**"truth"**	**results** Novartis (Basel, Ch)	NIBSC (UK)	PEI (Ger)
Beloc (BE)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**1**
	0.50	**1**	**0**	**0**	**1**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
Binotal (BI)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**1**	**1**	**1**
	0.50	**1**	**1**	**1**	**1**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
Ethanol 13% (ET)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**1**	**1**	**1**
	0.50	**1**	**1**	**1**	**1**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
Fenistil (FE)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**1**
	0.50	**1**	**0**	**0**	**1**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
Glucose 5% (GL)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**1**
	0.50	**1**	**1**	**1**	**1**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
"Drug A" 0.9% NaCl (LO)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**1**
	0.50	**1**	**1**	**1**	**1**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
MCP (ME)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**1**
	0.50	**1**	**1**	**1**	**1**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**

drug (code)	spike EU/ml	"truth"	results Novartis (Basel, Ch)	NIBSC (UK)	PEI (Ger)
"Drug B" 0.9% NaCl (MO)	0.00	0	0	0	0
	0.25	0	0	0	1
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Orasthin (OR)	0.00	0	0	0	0
	0.25	0	0	0	1
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Sostril (SO)	0.00	0	0	0	0
	0.25	0	0	0	1
	0.50	1	0	1	1
	0.50	1	0	1	1
	1.00	1	1	1	1

"0"denotes "non-pyrogenic"; "1" denotes "pyrogenic".

Table 13.4.2: 2x2 contingency table. PM applied to the PBMC-CRYO/IL-6 test result of 10 different substances assessed in three different laboratories. Results of each laboratory separately (DL, NL1 and NL2= Novartis, NIBSC and PEI respectively).

Results DL		True status of samples +	-	Total
PM	+	26	2	28
	-	4	18	22
Total		30	20	**50**

Results NL1		True status of samples +	-	Total
PM	+	28	2	30
	-	2	18	20
Total		30	20	**50**

Results NL2		True status of samples +	-	Total
PM	+	30	10	40
	-	0	10	10
Total		30	20	**50**

Table 13.4.2: 2x2 contingency table. PM applied to the PBMC-CRYO/IL-6 test result of 10 different substances assessed in three different laboratories (from table 13.4.1).

	True status of samples +	-	**Total**
PM +	84	14	**98**
-	6	46	**52**
Total	**90**	**60**	**150**

The overall specificity of the PBMC-CRYO/IL-6 assay is 76.7% (46/(46+14)*100). The overall sensitivity equals 93.3% (84/(84+6) *100%). Within the laboratories, specificity varied from 50% in one participating laboratory up to 90% in the other two laboratories. The sensitivity varied from 86.7% up to 93.3% and 100%.

Table 13.4.3: Specificity and sensitivity of the PBMC/IL-6 assay as determined from table 13.4.2.

	N total	**N correctly identified**	**proportion**	**95% CI lower limit**	**95% CI upper limit**
Specificity (Sp)	60	46	76.7%	64.7%	87.5%
Sensitivity (Se)	90	84	93.3%	85.6%	97.4%

13.5 Test Method Reliability (Reproducibility)

The interlaboratory reproducibility of the PBMC-CRYO/IL-6 method was assessed from the results of the catch-up validation testing 10 substances spiked with 5 separate spikes. The reproducibility varied from 76% to 80% and 96% between two laboratories. All three laboratories found the same result for 38 out of 50 samples.

Table 13.5.1: Interlaboratory reproducibility: Assessed by testing of 10 substances, spiked 5 times. One run of 50 samples by three different laboratories.

Laboratories	**Interlaboratory Reproducibility**	**Number of equal predictions**
DL - NL1	96%	48 / 50
DL - NL2	76%	38 / 50
NL1 – NL2	80%	40 / 50
Mean	84%	
same result in all labs	76%	38 / 50

DL =Novartis; NL1 = NIBSC; NL2 = PEI

13.6 Conclusion

It appears that accuracy and reproducibility of the test method do not decrease when cryopreserved cells replace fresh PBMCs. The sensitivity, which is well above 90%, is considered quite satisfactory for both methods. The somewhat lower results for specificity can fully explained by several false positives in just one of the laboratories. In the other two laboratories the specificity is still 90%, with a corresponding reproducibility of 96%.

14 Supporting Materials (Appendices)

14.1 Standard operating procedure (SOP) of the proposed method

Provide the complete, detailed protocol for the proposed test method.
Appendix A includes the PBMC/IL-6 test method protocol using fresh blood cells as used throughout the studies described in section 5 of the current BRD.
In addition, Appendix A includes the PBMC-CRYO/IL-6 test method protocol using cryopreserved blood cells. The protocol was used in catch-up validation study described in Section 13 of this BRD.
The trial plans of both studies are also included in Appendix A.

14.2 Standard operating Procedure (SOP) of the reference method

Provide the detailed protocol(s) used to generate reference data for this submission and any protocols used to generate validation data that differ from the proposed protocol.

14.3 Publications

Provide copies of all relevant publications, including those containing data from the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method with established performance standards.

Not all of the publications cited in the BRD (see section 12) are included as hardcopies in Appendix B, e.g. publications on statistical methods are not given.

Remark: The same set of hardcopies was included into Appendix B of all of the 5 submitted BRDs. Therefore some of the publications in Appendix B might not be referenced in the current BRD nor included in the list of the publications in section 12. Three general publications were added, which are not cited in any BRD but might be useful as background information to the validation study: the ECVAM publications on validation (Balls et al, 1995; Curren et al, 1995; Hartung et al, 2004). Several publications were included, which either give more background information on the human fever reaction or report on specific studies using *in vitro* pyrogen tests.

List of hard copies

Andrade SS, Silveira RL, Schmidt CA, Junior LB, Dalmora SL. (2003) Comparative evaluation of the human whole blood and human peripheral blood monocyte tests for pyrogens. Int J Pharm. Oct 20;265(1-2):115-24.

Balls et al. (1995) Practical aspects of the validation of toxicity test procedures. ECVAM Workshop Report 5. ATLA 23, 129-147.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Carlin G, Viitanen E. (2003) In vitro pyrogenicity of a multivalent vaccine: Infanrix. PHARMEUROPA, 15(3), 418-423.

Curren et al (1995) The Role of prevalidation in the development, validation and acceptance of Alternative Methods. ECVAM Prevalidation Task Force Report 1. ATLA 23, 211-217.

De Groote, D., Zangerle, P.F., Gevaert, Y., Fassotte, M.F., Beguin, Y., Noizat-Pirenne, F., Pirenne, J., Gathy, R., Lopez, M., Dehart I., (1992). Direct stimulation of cytokines (IL-1 beta, TNF-alpha, IL-6, IL-2, IFN-gamma and GM-CSF) in whole blood. I. Comparison with isolated PBMC stimulation. Cytokine 4, 239.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Fennrich S, Wendel A and Hartung T. (1999). New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 16:146-149

Fennrich S, Fischer M, Hartung T, Lexa P, Montag-Lessing T, Sonntag H-G, Weigandt M und Wendel A. (1999). Detection of endotoxins and other pyrogens using human whole blood. Dev. Biol. Standards 101:131-139

Gaines Das et al (2004). Monocyte activation test for pro-inflammatory and pyrogenic contaminants of parenteral drugs: test design and data analysis. J. of Immunological Methods 288, 165-177.

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hartung et al (2004) A modular approach to the ECVAM principles on test validity. ATLA 32, 467-472.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005). Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Peterbauer A, Schindler S, et al (2005). International validation of novel pyrogen tests based on human monocytoid cells. J Immunol Methods 298: 161-173.

Jahnke M, Weigand, Sonntag H-G. (2000). Comparative testing for pyrogens in parenteral drugs using the human whole blood pyrogen test, the rabbit in vivo pyrogen test and the LAL test. European Journal of Parenteral Science 5(2):39-44

Nakagawa Y, Maeda H, Murai T. (2002) Evaluation of the in vitro pyrogen test system based on proinflammatory cytokine release from human monocytes: comparison with a human whole blood culture test system and with the rabbit pyrogen test. Clin Diagn Lab Immunol. May;9(3):588-97.

Pasare, C. and Medzhitov, R. (2003). Toll pathway-dependent blockade of CD4+CD25+ T cell-mediated suppression by dendritic cells. Science 299: 1033-1036.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1998) The detection of pyrogens in blood products using an ex vivo whole blood culture assay. J Immunoassay. May-Aug;19(2-3):95-111.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1999) Differentiation between endotoxin and non-endotoxin pyrogens in human albumin solutions using an ex vivo whole blood culture assay. J Immunoassay. Feb-May; 20(1-2):79-89.

Poole S, Thorpe R, Meager A, Hubbard AJ, Gearing AJ (1988). Detection of pyrogen by cytokine release. Lancet 8577, 130.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, November 1989: 17-18.

Poole S, Musset MV (1989). The International Standard for Endotoxin: evaluation in an international collaborative study. J Biol Stand 17: 161-171.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' in vitro test for pyrogens. J Immunol Methods 274: 209-220.

Ray A, Redhead K, Selkirk S, Polle S (1991) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211-218.

Reddi, K., Nair, S.P. et al (1996). Surface-associated material from the bacterium Actinobacillus actinomycetemcomitans contains a peptide which, in contrast to lipopolysaccharide, directly stimulates fibroblast interleukin-6 gene transcription. Eur. J. Biochem. 236: 871-876.

Schindler S, Bristow A, Cartmell T, Hartung T and Fennrich S. (2003). Comparison of the reactivity of human and rabbit blood towards pyrogenic stimuli. ALTEX 20:59-63.

Schindler et al (2004) Cryopreservation of human whole blood for pyrogenicity testing. Journal of Immunological Methods 294, 89–100

Schins RP, van Hartingsveldt B, Borm PJ. Ex vivo cytokine release from whole blood. A routine method for health effect screening. Exp Toxicol Pathol. 1996 Nov; 48(6):494-6.

Spreitzer I, Fischer M, hartzsch K, Lüderitz-Püschel U and Montag T. (2000). Comparative Study of Rabbit Pyrogen Test and Human whole Blood Assay on Human Serum Albumin. ALTEX 19 (Suppl. 1):73-75

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

14.4 Original data

Include all available non-transformed original data for both the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method with established performance standards.

NOTE: The original data of the ELISA-plate reader were collected by S.Hoffman and ECVAM. These are available on the CD which goes with the BRD.

14.5 Performance standards

If appropriate performance standards for the proposed test method do not exist, performance standards for consideration by NICEATM and ICCVAM may be proposed. Examples of established performance standards can be located on the ICCVAM / NICEATM web site at http://iccvam.niehs.nih.gov.

BRD: PBMC/IL-6 March, 2006

APPENDIX A

Trial plan "Comparison And Validation Of Novel Pyrogen Tests Based On The Human Fever Reaction" Acronym: Human (e) Pyrogen Test

BRD: PBMC/IL-6 March, 2006

Detailed protocol PBMC/IL-6: "In Vitro Pyrogen Test Using PBMC (SP+PB var. Novartis) 03 10 02

Trial plan "Catch-up Validation of Novel Pyrogen Tests Based on the Human Fever Reaction"

Detailed protocol PBMC/IL-6 CRYO: In vitro Pyrogen Test Using Freshly Taken or Cryopreserved PBMC (SP+PB var. Novartis)

BRD: PBMC/IL-6 March, 2006

APPENDIX B

Not all of the publications cited in the BRD (see section 12) are included as hardcopies in Appendix B, e.g. publications on statistical methods are not given.

Remark: The same set of hardcopies was included into Appendix B of all of the 5 submitted BRDs. Therefore some of the publications in Appendix B might not be referenced in the current BRD nor included in the list of the publications in section 12. Three general publications were added, which are not cited in any BRD but might be useful as background information to the validation study: the ECVAM publications on validation (Balls et al, 1995; Curren et al, 1995; Hartung et al, 2004). Several publications were included, which either give more background information on the human fever reaction or report on specific studies using *in vitro* pyrogen tests.

List of hard copies

Andrade SS, Silveira RL, Schmidt CA, Junior LB, Dalmora SL. (2003) Comparative evaluation of the human whole blood and human peripheral blood monocyte tests for pyrogens. Int J Pharm. Oct 20;265(1-2):115-24.

Balls et al. (1995) Practical aspects of the validation of toxicity test procedures. ECVAM Workshop Report 5. ATLA 23, 129-147.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Carlin G, Viitanen E. (2003) In vitro pyrogenicity of a multivalent vaccine: Infanrix. PHARMEUROPA, 15(3), 418-423.

Curren et al (1995) The Role of prevalidation in the development, validation and acceptance of Alternative Methods. ECVAM Prevalidation Task Force Report 1. ATLA 23, 211-217.

De Groote, D., Zangerle, P.F., Gevaert, Y., Fassotte, M.F., Beguin, Y., Noizat-Pirenne, F., Pirenne, J., Gathy, R., Lopez, M., Dehart I., (1992). Direct stimulation of cytokines (IL-1 beta, TNF-alpha, IL-6, IL-2, IFN-gamma and GM-CSF) in whole blood. I. Comparison with isolated PBMC stimulation. Cytokine 4, 239.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Fennrich S, Wendel A and Hartung T. (1999). New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 16:146-149

Fennrich S, Fischer M, Hartung T, Lexa P, Montag-Lessing T, Sonntag H-G, Weigandt M und Wendel A. (1999). Detection of endotoxins and other pyrogens using human whole blood. Dev. Biol. Standards 101:131-139

Gaines Das et al (2004). Monocyte activation test for pro-inflammatory and pyrogenic contaminants of parenteral drugs: test design and data analysis. J. of Immunological Methods 288, 165-177.

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hartung et al (2004) A modular approach to the ECVAM principles on test validity. ATLA 32, 467-472.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005). Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Peterbauer A, Schindler S, et al (2005). International validation of novel pyrogen tests based on human monocytoid cells. J Immunol Methods 298: 161-173.

Jahnke M, Weigand, Sonntag H-G. (2000). Comparative testing for pyrogens in parenteral drugs using the human whole blood pyrogen test, the rabbit in vivo pyrogen test and the LAL test. European Journal of Parenteral Science 5(2):39-44

Nakagawa Y, Maeda H, Murai T. (2002) Evaluation of the in vitro pyrogen test system based on proinflammatory cytokine release from human monocytes: comparison with a human whole blood culture test system and with the rabbit pyrogen test. Clin Diagn Lab Immunol. May;9(3):588-97.

Pasare, C. and Medzhitov, R. (2003). Toll pathway-dependent blockade of CD4+CD25+ T cell-mediated suppression by dendritic cells. Science 299: 1033-1036.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1998) The detection of pyrogens in blood products using an ex vivo whole blood culture assay. J Immunoassay. May-Aug;19(2-3):95-111.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1999) Differentiation between endotoxin and non-endotoxin pyrogens in human albumin solutions using an ex vivo whole blood culture assay. J Immunoassay. Feb-May; 20(1-2):79-89.

Poole S, Thorpe R, Meager A, Hubbard AJ, Gearing AJ (1988). Detection of pyrogen by cytokine release. Lancet 8577, 130.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of

the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, November 1989: 17-18.

Poole S, Musset MV (1989). The International Standard for Endotoxin: evaluation in an international collaborative study. J Biol Stand 17: 161-171.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' in vitro test for pyrogens. J Immunol Methods 274: 209-220.

Ray A, Redhead K, Selkirk S, Polle S (1991) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211-218.

Reddi, K., Nair, S.P. et al (1996). Surface-associated material from the bacterium Actinobacillus actinomycetemcomitans contains a peptide which, in contrast to lipopolysaccharide, directly stimulates fibroblast interleukin-6 gene transcription. Eur. J. Biochem. 236: 871-876.

Schindler S, Bristow A, Cartmell T, Hartung T and Fennrich S. (2003). Comparison of the reactivity of human and rabbit blood towards pyrogenic stimuli. ALTEX 20:59-63.

Schindler et al (2004) Cryopreservation of human whole blood for pyrogenicity testing. Journal of Immunological Methods 294, 89–100

Schins RP, van Hartingsveldt B, Borm PJ. Ex vivo cytokine release from whole blood. A routine method for health effect screening. Exp Toxicol Pathol. 1996 Nov; 48(6):494-6.

Spreitzer I, Fischer M, hartzsch K, Lüderitz-Püschel U and Montag T. (2000). Comparative Study of Rabbit Pyrogen Test and Human whole Blood Assay on Human Serum Albumin. ALTEX 19 (Suppl. 1):73-75

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

BRD: PBMC/IL-6 March, 2006

APPENDIX C

List of abbreviations and definitions

Accuracy	The ability of a test system to provide a test result close to the accepted reference value for a defined property.
BET	The bacterial endotoxin test is used to detect or quantify endotoxins of gram-negative bacterial origin using amoebycte lysate from horseshoe crab (*Limulus polyphemus* or *Tachypleus tridentatus*
BRD	Background Review Document
CRYO WB/IL-1	Whole blood assay (using cryopreserved blood) with IL-1 as endpoint
CV	coefficient of variation
DL	Developing laboratory = laboratory which developed the method or the most experienced laboratory
ELC	Endotoxin limit concentration; maximum quantity of endotoxin allowed in given parenterals according to European Pharmacopoeia
Endotoxins	Endotoxins are a group of chemically similar cell-wall structures of Gram-negative bacteria, i.e. lipopolysaccharides
ELISA	Enzyme linked immunosorbent assay
EU/ml	European Units per ml
IL-1	interleukin 1
IL-6	interleukin 6
Intralaboratory reproducibility	A determination of the extent that qualified people within the same laboratory can independently and successfully replicate results using a specific protocol at different times.
Interlaboratory reproducibility	A measure of the extent to which different qualified laboratories, using the same protocol and testing the same substances, can produce qualitatively and quantitatively similar results. Inter-laboratory reproducibility is also referred to as between-laboratory reproducibility.
KN	University of Konstanz (Konstanz, Germany), developing laboratory WB/IL-1 and CRYO WB/IL-1
LPS	lipopolysaccharides
MM6	MONO MAC-6 cell line

MM6/IL-6	In vitro pyrogen test using MM6 cell line and IL-6 release as an endpoint
MVD	Maximum valid dilution; the MVD is the quotient of the ELC and the detection limit
NIBSC	National Institute for Biological Standards and Control (London, UK), developing laboratory for WB/IL-6
NL	naïve laboratory = laboratory with non or minor experience with the method
NPC	negative product control (clean, released lot of the nominated product under test)
Novartis	Novartis (Basel, Switzerland), developing laboratory PBMC/IL-6
OD	optical density
PBMC	Peripheral blood mononuclear cells
PBMC/IL-6	In vitro pyrogen test using fresh peripheral blood mononuclear cells and IL-6 release as endpoint
PBMC-CRYO/IL-6	In vitro pyrogen test using cryopreserved peripheral blood mononuclear cells and IL-6 release as endpoint
PEI	Paul-Ehrlich Institut (Langen, Germany), participating laboratory
PM	prediction model = is an explicit decision-making rule for converting the results of the in vitro method into a prediction of in vivo hazard
PPC	positive product control (product under test spiked with 0.5 EU/ml of WHO-LPS (code 94/580)
Prevalidation study	A prevalidation study is a small-scale inter-laboratory study, carried out to ensure that the protocol of a test method is sufficiently optimised and standardised for inclusion in a formal validation study. According to the ECVAM principles, the prevalidation study is divided into three phases: protocol refinement, protocol transfer and protocol performance (Curren et al, ATLA 23, 211-217).
Pyrogens	fever-causing materials
Pyrogens, endogenous	endogenous pyrogens are messenger substances released by blood cells reacting to pyrogenic materials; e.g. IL-1, IL-6, TNF-α, prostaglandin E_2
Pyrogens, exogenous	exogenous pyrogens derive from bacteria, viruses, fungi or from the host himself
Reliability	Measures of the extent to which a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is

	assessed by calculating intra- and interlaboratory reproducibility and intra-laboratory repeatability.
Relevance	Relevance of a test method describes whether it is meaningful and useful for a particular purpose. It is the extent to which the measurement result and uncertainty can accurately be interpreted as reflecting or predicting the biological effect of interest.
Repeatibility	Repeatability describes the closeness of agreement between test results obtained within a single laboratory when the procedure is performed independently under repeatability conditions, i.e. in a set of conditions including the same measurement procedure, same operator, same measuring system, same operating conditions and same location, and replicated measurements over a short period of time.
RIVM	National Institute of Public Health and the Environment (Bilthoven, The Netherlands), developing laboratory MM6/IL-6 method
Sensitivity	Sensitivity is the proportion of all positive/active substances that are correctly classified by a test method.
Specificity	Specificity is proportion of all negative/inactive substances that are correctly classified by a test method.
TMB	chromogenic substrate 3,3´,5,5´ -tetramethylbenzidine
TNF-α	tumour necrosis factor-α
USP	US Pharmacopoeia
Validation	Validation is the process by which the reliability and relevance of a procedure are established for a specific purpose
Validation study	A validation study is a large-scale interlaboratory study, designed to assess the reliability and relevance of an optimised method for a particular purpose
WB/IL-1	Whole blood assay (using fresh blood) with IL-1 release as endpoint
WB/IL-6	Whole blood assay (using fresh blood) with IL-6 release as endpoint
WHO	World Health Organization

[This Page Intentionally Left Blank]

Validation of Biomedical Testing Methods

In vitro Pyrogen Test Using Freshly Taken or Cryopreserved PBMC

(SP+PB var. Novartis)

Standard Operating Procedure

Only the GLP/QA Unit is allowed to make copies of this document. Extra copies can be obtained on request from the GLP/QA Unit. Quality Documents are valid only if they are signed by the GLP/QA Unit and provided with a valid copy number.	**Copy number**

Date of Circulation: *July 02, 2004*
File Identifier: *SOP-PBMC2.v01*

JOINT RESEARCH CENTRE
EUROPEAN COMMISSION
Institute for Health and Consumer Protection
ECVAM

Standard Operating Procedure

In vitro Pyrogen Test Using Freshly Taken or Cryopreserved PBMC (SP+PB var. Novartis)

Drafted by:	Name Date Signature	Peter Bruegger
Reviewed by:	Name Date Signature	Stephen Poole
Approved by:	Name Date Signature	
Issued by:	Name Date Signature	

*Owner/Trainer: Signature: Date:

PAGE OF CHANGES

Date of change/ Date of draft:	Version-number:	Changed page(s):	Summary of the change(s):	Changed by/Sign.:

1 INTRODUCTION 2
2 PURPOSE 2
3 SCOPE / LIMITATIONS 2
4 METHOD OUTLINE 2
5 DEFINITIONS / ABBREVIATIONS 2
6 MATERIALS 2
6.1 Buffers and Reagents for Novartis IL-6 ELISA: 2
7 METHODS 2
7.1 Coating of IL-6 ELISA plates 2
7.2 Preparation of samples for assay (*) 2
7.3 Collection of human blood 2
7.4 Isolation of PBMCs(*) 2
7.5 Washing PBMCs(*) 2
7.6 Procedure for cryopreserving and thawing PBMCs(*) 2
7.7 Equilibration of reagents for cell culture 2
7.8 Preparation of the LPS standard curve (*) 2
7.9 Cell culture (*) 2
7.10 Detection of IL-6 in the supernatant medium by ELISA 2
7.10.1 Preparation of IL-6 standard curve 2
7.10.2 Addition of standards and samples 2
7.10.3 Addition of substrate solution and measuring 2
8 DATA ANALYSIS, PREDICTION MODEL AND RELATED ERRORS 2
8.1 Acceptance Criteria 2
8.2 Interference Test 2
Interference with the cell system 2
8.3 Interference with the ELISA system 2
8.4 Prediction model 2
9 HEALTH SAFETY AND ENVIRONMENT 2
10 REFERENCES 2

1 INTRODUCTION

Parenteral pharmaceutical products must be shown to be free from pyrogenic (fever-inducing) contamination. While a pyrogen may in general be defined as any substance that causes fever, the pyrogens that almost invariably contaminate parenteral pharmaceuticals are bacterial endotoxins (lipopolysaccharides, LPS) from Gram-negative bacteria (Mascoli and Weary, 1979a, 1979b). There are two Pharmacopoeial tests for pyrogenic contamination: the rabbit pyrogen test and the Limulus amoebocyte lysate (LAL) test. The rabbit pyrogen test, which detects LPS and other pyrogens, involves measuring the rise in body temperature evoked in rabbits by the intravenous injection of a sterile solution of the substance to be examined. In contrast, the LAL test detects only LPS: it is described in Pharmacopoeias as the bacterial endotoxins test (BET). The principle of the LAL-test is that LPS causes extracellular coagulation of the blood (haemolymph) of the horseshoe crab, *Limulus polyphemus.* (Levin & Bang, 1964). Although the LAL test is gradually superseding the rabbit pyrogen test, hundreds of thousands of rabbit pyrogen tests are still carried out each year around the world, largely on products which cannot, for one reason or another, be tested in the LAL test. While proving generally reliable, both the rabbit pyrogen test and LAL test have shortcomings. The rabbit pyrogen test uses experimental animals, is costly and is not quantitative. The LAL test gives false negatives with certain products, can overestimate the pyrogen content of other products and does not detect pyrogens other than bacterial endotoxin (LPS), such as Gram-positive exotoxins, viruses and fungi (Dinarello et al., 1984; Poole et al., 1988; Ray et al., 1990; Taktak et al., 1991; Fennrich et al., 1999).
The basis of the rabbit pyrogen test is the in vivo stimulation by exogenous pyrogens (usually LPS) of rabbit peripheral blood monocytes to produce the endogenous pyrogens that cause fever. The endogenous pyrogens are pyrogenic cytokines such as tumour necrosis factorα (TNFα), interleukin-1 (IL-1α and IL-1β, two separate gene products), IL-6 and IL-8 (Dinarello et al., 1999). In view of the shortcomings of the rabbit pyrogen test and the LAL test, in vitro pyrogen tests that utilise the exquisite sensitivity to exogenous pyrogen of monocytes have been proposed. In such tests, products are incubated with human peripheral blood monocytes (or mononuclear cells, PBMC, or leukocytes) and the conditioned media assayed for pyrogenic cytokines (Duff & Atkins, 1982; Dinarello et al., 1984; Poole et al., 1988; Poole, 1989; Hansen and Christensen, 1990; Taktak et al., 1991; Bleeker et al., 1994).

2 PURPOSE

By quantifying the amounts of cytokines released by PBMC stimulated with the USP reference preparation for endotoxin or the international standard (IS) for endotoxin (LPS), various non-endotoxin pyrogens and with medicinal products spiked with endotoxin, it is the objective that an in vitro pyrogen test be developed that will serve as a replacement for the rabbit pyrogen test.

3 SCOPE / LIMITATIONS

The method described below is for the evaluation of in vitro PBMC/cytokine release tests using IL-6 as the cytokine that serves as the readout (measured variable). It is not a 'finalised' test system for the testing of medicinal products. Especially the prediction model has to be developed cautiously.

The method may be applied only to preparations that have been validated with the method, i.e. shown not to interfere by causing inhibition or enhancement of LPS (STD)-induced cytokine production.

4 METHOD OUTLINE

Freshly taken human whole blood is heparinised, diluted with phosphate buffered saline (PBS) and the PBMC's isolated. For storage and shipping purposes PBMCs may be cryopreserved. Freshly taken or thawed PBMCs are then stimulated for 16-24h with the USP reference preparation for endotoxin, the international standard for endotoxin, (LPS) or samples of related materials, e.g. other endotoxins, non-endotoxin pyrogens and medicinal products unspiked and spiked with endotoxin. Following this stimulation, the concentrations of the cytokine in the PBMC-conditioned medium are quantified using a specific ELISA for IL-6 (which is calibrated in terms of the appropriate international standard). The construction of dose-response curves for endotoxin (LPS) versus concentrations of released cytokines permits the estimation of the endotoxic/cytokine-releasing activity contained in the samples.

5 DEFINITIONS / ABBREVIATIONS

Ab	Antibody
BSA	Bovine serum albumin
CO_2	Carbon dioxide
°C	Degrees Celsius (Centigrade)
DMSO	Dimethylsulfoxide
D-R	Dose-response
E. coli	Escherichia coli
ELISA	Enzyme-linked immunosorbent assay
EP	European Pharmacopoeia
EU	Endotoxin units
FDA	Food and Drug Administration (USA)
g	Gram
h	Hour
HIFCS	Heat-inactivated (+56°C for 30 min) foetal calf serum
H_2O_2	Hydrogen peroxide
H_2SO_4	Sulphuric acid
IL	Interleukin
IS	International standard
IU	International unit
l	Litre
KOH	Potassium hydroxide
LAL	Limulus amoebocyte lysate
LPS	Lipopolysaccharide
M	Molar
MAb	Monoclonal antibody
mg	Milligram
min	Minute
ml	Millilitre
mM	Millimolar
µg	Microgram
µl	Microlitre
NaCl	Sodium chloride
NaOH	Sodium hydroxide
$NaHCO_3$	Sodium hydrogen carbonate
NaH_2PO_4	Sodium di-hydrogen phosphate
Na_2HPO_4	di-Sodium hydrogen orthophosphate
nm	Nanometre
OD	Optical density
PBMC	Peripheral blood mononuclear cells
PBS	Dulbecco's Phosphate Buffered Saline
PF	Pyrogen-free (items purchased as sterile and pyrogen-free or baked at 250° for 30-60 min.
POD	Horseradish peroxidase conjugate
R	Endotoxin standard
rpm	Rounds per minute
RPMI	RPMI 1640 cell culture medium
RPMI-C	RPMI 1640 cell culture medium + human AB serum at a final concentration of 2% v/v
RSE	Reference Standard Endotoxin
RT	Room temperature
TMB	Tetramethyl benzidine
S	Test sample
STD	Standard

UNK	Unknown
USP	United States Pharmacopoeia
x g	x gravity

6 MATERIALS

Sterile, pyrogen-free Dulbecco's phosphate buffered saline w/o Ca^{++} and Mg^{++} (Life Technologies)
Polyoxyethylene-sorbitan monolaurate (TWEEN 20), cell culture grade, (Sigma, P-2287)
Hydrochloric acid, 0.1M, sterile filtered (Sigma, H-9892)
Sodium hydroxide (reagent grade)
1M H_2SO_4 (Merck)
Mouse monoclonal anti-IL-6 antibody from clone 16
Horseradish peroxydase conjugated sheep polyclonal anti-IL-6 antibody
3,3',5,5'-Tetramethyl benzidine (e.g. Fluka Cat. No. 87748)
Acetone (reagent grade)
Ethanol (reagent grade)
Phenole (e.g. Merck Cat. No. 100206)
Potassium hydroxide (reagent grade)
Sodium dihydrogen phosphate (e.g. Merck Cat. No. 106346)
Disodium hydrogen phosphate (e.g. Merck Cat. No. 106580)
Tris (hydroxymethyl) aminomethane (e.g. Fluka Cat. No. 93352)
Kathon MW/WT, Christ Chemie AG, Reinach, Switzerland
Albumin from bovine serum (e.g. Fluka Cat. No. 05480)
Citric acid monohydrate e.g. Fluka Cat. No. 27490)
Human AB serum (Sigma)
Trypan blue stain (Sigma)
USP Reference Standard Endotoxin [EC6 lot G], identical to the WHO international standard for bacterial endotoxin (LPS, vial code 94/580)
Fragmin (Dalteparin, 10000 IU/ml, Pharmacia)
RPMI 1640 medium (Life Technologies™, Paisley, Scotland)
L-Glutamine 200mM (Life Technologies™, Paisley, Scotland)
Penicillin/Streptomycin solution (Seromed Cat. No. A2213)
Lymphoprep (Nycomed, Oslo, Norway)
Dimethylsulfoxide, LAL tested, to be free of detectable endotoxins.
Nunc-Immuno 96-well plate MaxiSorp (F96, Life Technologies™, Paisley, Scotland)
Falcon Microtest tissue culture plate, 96-well (353072, Beckton Dickinson Labware)
Falcon serological pipettes (5ml,10ml, 25ml, Beckton Dickinson Labware)
Centrifuge tubes (Falcon 2070 Blue Max™)
Polypropylene conical tubes (Falcon 2069 Blue Max™)
Eppendorf Biopur Tips 100ul & 1000ul (Eppendorf-Netheler-Hinz-Gmbh,Germany)
0.22 μm sterile filters (MilliPak 60, Millipore)
Eppendorf® volumetric pipettes
Pyrogen-free reservoir liner, 12-well (PMP-380-507L, Fisher, UK)

All other consumables are purchased as sterile and pyrogen-free and other reagents are pro analysis grade.

6.1 Buffers and Reagents for Novartis IL-6 ELISA:

Coating Buffer

Dissolve
5.0 g of sodium dihydrogen phosphate, (e.g. Merck Art. No. 106346), and
2.9 g of disodium hydrogen phosphate, (e.g.Merck Art. No. 106580),
in 400 ml of distilled water.

Use 1 N NaOH to adjust the pH to 7.5, and make up to 500 ml with distilled water.

Remains stable for 6 months at 2 -8°C.

Blocking Buffer

Tris(hydroxymethyl)aminomethane, (e.g.Fluka Art. No. 93352)	12.1 g
Dissolve in distilled water	400 ml
Kathon MW/WT, Christ Chemie AG, Switzerland	0.1 ml

Use 4 M HCl to adjust the pH to 7.5.

Albumin from bovine serum, (e.g.Fluka Art. No. 05480)	5.0 g
Add distilled water to make up to 500 ml.	

Remains stable for 6 months at 2 - 8°C.

Stopping Solution

Distilled water	500 ml
H_2SO_4	26.6 ml

Wash Solution

Demineralised water	2000 ml
Tween-20	1 ml

Interleukin-6 Standard

Use IL-6 from human lymphocytes (e.g. Boehringer Mannheim, Cat. No. 1299972). 1 vial contains 200 000 units of natural human IL-6.
Dilute the contents of a vial (1 ml) of human IL-6 with 4 ml of RPMI-C, and freeze aliquots of 100 µl at about – 80 °C.
Before the first use of a new batch, this IL-6 used as the assay calibrant has to be calibrated against the IS for IL-6 (89/548).

Mouse Anti-Human IL-6 Monoclonal Antibody

(*Reagent 1)*
derived from clone 16.

Detection Antibody (POD)

(Reagent 2)
Sheep anti-human IL-6 antibody, horseradish peroxidase conjugated. Remains stable for at least 6 months at 2-8°C.

Human IL-6 Standard Solution

(Reagent 3)
Thaw 1 frozen aliquot of interleukin-6 standard (4000 units in 100 µl) and dilute with 900 µl of RPMI-C.
Add 100 µl of this dilution to 900 µl of RPMI-C = 400 units/ml (standard initial concentration, equivalent to 4000 pg/ml).
Prepare the solution shortly before use; do not store.

Dilution Buffer

(*Reagent 4)*
Prepare the dilution buffer as follows:

Tris(hydroxymethyl)aminomethane, (e.g.Fluka Art. No. 93352)	2.1 g
Distilled water	400 ml
Kathon MW/WT, Christ AG, Switzerland	0.1 ml
Phenol, (e.g.Merck Art. No. 100206)	0.5 g
Heat-inactivated (30 minutes/56 °C) fetal bovine serum	25 ml

Mix to dissolve the substances, then adjust the pH to 7.5 with 4 M HCl. Make up to 500 ml with distilled water.

Remains stable for at least 6 months at 2 - 8 °C.
In the absence of the stabilizers Kathon and phenol the stability is only 1 day.

TMB Solution

(Reagent 5)

Prepare the TMB solution as follows:

3,3',5,5'Tetramethylbenzidine (e.g.Fluka Art. No. 87748)	240 mg
Reagent-grade acetone	5 ml

Dissolve, then add

Reagent-grade ethanol	45 ml
Perhydrol (30 % H_2O_2), (e.g.Merck Art. No. 107209)	0.300 ml

Remains stable for at least 6 months at 15 - 25 °C when sealed and protected from light.

Substrate Buffer

(Reagent 6)

Reagent-grade citric acid monohydrate (e.g.Fluka Art. No. 27490)	6.3 g
Distilled water	800 ml

Mix to dissolve, then adjust the pH to 4.1 by adding 4 M KOH.
Make up to 1000 ml with distilled water and add 0.2 ml of Kathon MW/WT

Remains stable for about 6 months at 15 - 25 °C.
In the absence of the Kathon the stability is only 1 day.

Culture Medium (RPMI-C)

The preparation of the medium has to be carried out in a class 2 laminar flow sterile cabinet, using aseptic technique and reagents and consumables that are sterile and pyrogen-free.

RPMI 1640 medium	500 ml
Human serum AB	5 ml
L-Glutamine, 200 mM	5 ml
Penicillin/Streptomycin solution	10 ml

7 METHODS

Steps marked (*) are carried out in a class 2 laminar flow sterile cabinet, using aseptic technique and reagents and consumables that are sterile and pyrogen-free.

7.1 Coating of IL-6 ELISA plates

Dilute the coating anti-IL-6 antibody (Clone 16) with coating buffer to 2.5 µg/ml and swirl to mix, e.g. 1 mg of antibody in 400 ml of coating buffer. Add 200 µl to each well of a 96-well plate (Nunc-Immuno MaxiSorp F96)
Stack the microtitre plates and allow to stand in the dark at 15 -25°C.for 16-24 hours.

Aspirate and discard the coating solution. Wash the coated plate 3 times with demineralised water, and tap out onto cellulose.
Pipette 200 µl of blocking buffer into each of the wells to block the residual protein-binding capacity of the coated plates.

Seal the plates with adhesive film, and store in a humidified atmosphere at 2 - 8 °C (shelf life: two months).

7.2 Preparation of samples for assay (*)

Samples are tested at a dilution of 1 in 5, i.e. 50 µl of sample in a total culture volume of 250 µl. To test samples at dilutions greater than 1 in 5, pre-dilute samples before addition to the assay plate, e.g. to test a sample at a dilution of 1 in 10, pre-dilute the sample 1 in 2 with saline and add 50 µl of this diluted sample to the assay plate.

7.3 Collection of human blood

Qualification of blood donors: Blood donors are to describe themselves as being in good health, not suffering from any bacterial or viral infections (including colds and influenza). Blood donors are not to be taking drugs known to influence the production of cytokines. Also, See Section 8. (below) for the criteria for the rejection of data as having come from a test utilising blood from a donor not in good health.

***Procedure(*)*:** Dilute 1 ampoule of Fragmin (10000 IU in 1 ml) with 1 ml pyrogen-free distilled water. Draw back the piston of a 30 ml syringe by about 3 mm, then inject, using a syringe calibrated in µl, 50 µl of the diluted Fragmin via the luer lock adapter into the space between the syringe wall and the piston. Then attach a 19 mm, 21 gauge butterfly system to the 30 ml syringe. Alternatively a 40mm 21 gauge needle may be used.

Using this prepared (heparinised) syringe, draw 30 ml blood from the median cubital or cephalic vein of the left or right arm of a single donor. Mix the blood with the anticoagulant by tilting the syringe several times.

Afterwards remove the butterfly or needle and transfer 15 ml of the blood into each of two 50 ml sterile, pyrogen-free centrifuge tubes.

7.4 Isolation of PBMCs(*)

PBMCs are isolated using Lymphoprep. The procedure is a modification of the manufacturers' instructions that may permit cleaner separation of the PBMC. Start the procedure not later than 2 hours after blood withdrawal.

Procedure: Add 15 ml PBS to each 15 ml of heparinised whole blood to make 2 tubes of 30 ml of diluted blood. Break off the end of a 10 ml pipette to remove the cotton and then place the pipette carefully (point down) in the tube containing the diluted blood. Add 20 ml Lymphoprep to each of the two tubes, via the inserted pipette, to form a lower layer. Centrifuge at 340 x g for 45 min at room temperature (18 – 25 °C), with the centrifuge brake set to off/zero. After centrifugation, the PBMC form a white band at about the 25 ml graduation of the tube. Carefully draw off the uppermost 9ml of each of the supernatants from the same donor and pipette to a new 50 ml tube if it is intended to freeze the PBMCs, otherwise discard. Aspirate and discard the remaining supernatants from above the PBMCs. Take up the PBMCs with a 10 ml pipette and transfer to a new 50 ml centrifuge tube.

7.5 Washing PBMCs(*)

To the isolated PBMCs add sufficient PBS to give a total volume of 50 ml and centrifuge at 340 x g for 15 min.

Aspirate the supernatant phase, and resuspend the sediment with 10 ml of PBS using a serological pipette (aspirate and expel several times, do not vortex). Make up to a total volume to 50 ml with PBS and centrifuge at 340 x g for 10 min. Aspirate the supernatant phase and resuspend the sediment with 15 ml of RPMI-C. Pool the resuspended sediments for each donor into one tube per donor and distribute the cells with a volume of RPMI-C equal to the initial volume of blood from which the PBMC were isolated. This suspension of PBMC in RPMI-C is used in the cell-culture.
The cells are to be cultured with endotoxin or samples within 4 hours of blood withdrawal.

7.6 Procedure for cryopreserving and thawing PBMCs(*)

Add 2 ml of endotoxin-free DMSO to 18 ml of separated supernatant from each donor to produce the cryo-protective solution and cool to 2 to 8°C. Proceed as under point 7.5 of "Washing PBMCs" but centrifuge only once. Discard the supernatant and resuspend the cell pellet with 6 ml of the chilled, homologous cryoprotective solution. Pool the cell suspensions from the same donor and make aliquots of 1.0 ml in screw cap cryotubes. One 1.0 ml aliquot contains the PBMC of 5 ml whole blood. Slowly cool down the aliquoted cell suspensions to –80°C using a styrofoam box to provide thermal insulation. For prolonged storage transfer the tubes after 72 hours to liquid nitrogen (-196°C).

Take two tubes (2 x 1.0 ml) of cryopreserved cells from one donor from the liquid nitrogen and submerge the tubes (but not their caps) immediately in a water bath at a temperature of 37°C. After thawing, pool the cell suspension in a 50ml centrifuge tube and add culture medium to give a total volume of 40 ml. Take care not to contaminate the contents with the water from the water bath. Centrifuge for 10 minutes at 340 x g, pour off the supernatant phase and resuspend the sedimented cells in 10 ml of culture medium (RPMI-C). If desired, a viability test can be performed. (In our experience the viability is above 95%.)

7.7 Equilibration of reagents for cell culture

Bring a vial of the LPS standard, the samples for assay and a bottle of RPMI-C to room temperature.

7.8 Preparation of the LPS standard curve (*)

Add 5 ml of LAL reagent water to the lyophilised contents of one vial of the current USP Reference Standard Endotoxin or the IS to produce a stock solution of 2000 EU(IU)/ml. Vortex for at least 30 min. The stock solution remains stable for 14 days if stored at 2 – 8°C.
Prepare the LPS standard curve by making serial dilutions in saline of the stock solution of endotoxin:

Label seven tubes, A – G. Add the volumes of saline to the tubes specified in table 1, below.

Take 200µl of endotoxin stock solution and add 1.8 ml of saline and vortex to make 2.0 ml of a 200 EU (IU)/ml solution of LPS = Solution S.

Table 1. Preparation of the LPS standard curve:

Tube	LPS added to tube	Saline	[LPS] in tube	→ *[LPS] in well*
A	100 µl of Solution S = 20 EU	900 µl	20 EU/ml	*Not for culture*
B	100 µl of Solution A = 2 EU	1900 µl	1 EU/ml	*0.2 EU/ml*
C	500 µl of Solution B = 0.5 EU	500 µl	0.5 EU/ml	*0.1 EU/ml*
D	500 µl of Solution C = 0.25 EU	500 µl	0.25 EU/ml	*0.05 EU/ml*
E	500 µl of Solution D = 0.125 EU	500 µl	0.125 EU/ml	*0.025 EU/ml*
F	500 µl of Solution E = 0.063 EU	500 µl	0.063 EU/ml	*0.0125 EU/ml*
G	None	1 ml	0 EU/ml	*0 EU/ml*

Vortex each of Solutions A - G after its preparation and then use each solution for the preparation of the subsequent dilution.
(Multiples of the above volumes may be used to generate larger volumes of Solutions A – G and an LPS standard calibrated against the USP RSE or the IS for endotoxin may be used.)

7.9 Cell culture (*)

Add 100 µl of RPMI-C to wells of columns 1 – 10 as in template 1, see below.

Add 50 µl of the test samples S1 – S14 to wells as in template 1, see below.

Add 50 µl of LPS standards to wells as in template 1, below.

Solution G into wells A3 – D3 (STD R0)
Solution F into wells A4 – D4 (STD R1)
Solution E into wells A5 – D5 (STD R2)
Solution D into wells A6 – D6 (STD R3)
Solution C into wells A7 – D7 (STD R4)
Solution B into wells A8 – D8 (STD R5)

(The above order of addition permits the same tip to be used for additions of all the standards.)

Gently swirl the solution of PBMC to reduce settling of the cells and to distribute the PBMC more evenly throughout the RPMI-C solution immediately before aliquots of PBMC are taken. Do not vortex.

Add 100 µl of PBMC to the 96-well plate (see template 1, below). Add the PBMC by row in the following sequence: A, E, B, F, C, G, D, H if a repeating pipette is used. Alternatively a multipipette may be used for these additions provided that the aliquots are added briskly to minimise the settling of cells.

Gently swirl the resulting cultures to mix the contents of the wells without cross-contaminating wells.
Incubate the cultures without vibration (to allow the cells to settle) at 37°C for 16 - 24h in an atmosphere of 5% CO_2 in humidified air.

Template 1: PBMC culture plate

	1	2	3	4	5	6	7	8	9	10	11	12
A	S1	S2	R0 0	R1 0.063	R2 0.125	R3 0.25	R4 0.5	R5 1	S3	S4	void	void
B	S1	S2	R0 0	R1 0.063	R2 0.125	R3 0.25	R4 0.5	R5 1	S3	S4	void	void
C	S1	S2	R0 0	R1 0.063	R2 0.125	R3 0.25	R4 0.5	R5 1	S3	S4	void	void
D	S1	S2	R0 0	R1 0.063	R2 0.125	R3 0.25	R4 0.5	R5 1	S3	S4	void	void
E	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	void	void
F	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	void	void
G	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	void	void
H	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	void	void

Key:
S1 – S14 = test samples #1 - #14
R0 – R5 = Reference standard for endotoxin , R0 = 0 EU/ml, R1 = 0.063 EU/ml, R2 = 0.125 EU/ml, R3 = 0.25 EU/ml, R4 = 0.5 EU/ml and R5 = 1 EU/ml (*The final concentrations are: 0.0125, 0.025, 0.05, 0,1 and 0,2UE/ml).*

7.10 Detection of IL-6 in the supernatant medium by ELISA

Immunoreactive IL-6 in aliquots of the tissue culture fluid is quantified using a validated ELISA, in which the IL-6 standard used as the assay calibrant is calibrated against the IS for IL-6 (89/548).

7.10.1 Preparation of IL-6 standard curve

Prepare the IL-6 standard dilutions as follows:
Fill each of 7 polystyrene tubes (12ml) with 500 µl of RPMI-C.
Add 900 µl of RPMI-C to one frozen aliquot (100µl) of IL-6 standard (=Reagent 3)
Dilute this concentration (Solution I) by transferring 500 µl to tube J (1 in 2). Further dilute by transferring 500 µl from this tube to the next, 500 µl from that tube to the following one, and so on, ending with the tube marked with O. Use the RPMI-C in tube P as a blank (see table 2).

Table 2. Preparation of the IL-6 standard curve from an aliquot of the IL-6 standard.

Sol.	IL-6 added	RPMI-C	IL-6 in tube
I	100 µl of frozen sol.= 4000 pg	900 µl	4000 pg/ml
J	500 µl of Solution I = 2000 pg	500 µl	2000 pg/ml
K	500 µl of Solution J = 1000 pg	500 µl	1000 pg/ml
L	500 µl of Solution K = 500 pg	500 µl	500 pg/ml
M	500 µl of Solution L = 250 pg	500 µl	250 pg/ml
N	500 µl of Solution M = 125 pg	500 µl	125 pg/ml
O	500 µl of Solution N = 62.5 pg	500 µl	62.5 pg/ml
P	None	500 µl	0 pg/ml

Vortex each of Solutions I – O after its preparation and then use each solution for the preparation of the subsequent dilution.
(*Multiples of the above volumes may be used to generate larger volumes of Solutions I – P).*

7.10.2 Addition of standards and samples

Volumes are sufficient for 4 microtiter plates.
Dilute detection antibody (Reagent 2) with dilution buffer (Reagent 4) (e.g. 1 in 500 [200 µl of Reagent 2 + 100 ml of Reagent 4]), and mix without causing foam to form Test each batch of detection antibody in separate experiments to determine the optimum dilution.

Just before using shake out the blocking buffer from the antibody-coated microtitre plates, place the plates on cellulose with the openings facing down, and tap.
At the end of the tissue culture incubation, transfer 50 µl of supernatant from each of the wells of columns 1 – 10 of the tissue culture plate to the corresponding wells on each of the cytokine ELISA plates – see template 1, above and template 2, below. A multichannel pipette may be used. Ensure that the well contents are mixed by aspirating and expelling 50 µl three times before transferring the liquid. (The wells in columns 11 and 12 are for the IL-6 standard curve – see, below).

Transfer 50 µl of each of the dilutions of the IL-6 standard, and of the blank, into 2 wells each (standard concentrations from 4000 - 62.5 pg/ml).
Add 200 µl of dilute detection antibody (e.g. 1 in 500) to each of the wells, seal the microtitre plates with adhesive film, and allow to stand for 2 - 3 hours at 20 - 25 °C. Use template 2 below.

Template 2: ELISA plate

	1	2	3	4	5	6	7	8	9	10	11	12
A	S1	S2	R0	R1	R2	R3	R4	R5	S3	S4	0	0
B	S1	S2	R0	R1	R2	R3	R4	R5	S3	S4	62.5	62.5
C	S1	S2	R0	R1	R2	R3	R4	R5	S3	S4	125	125
D	S1	S2	R0	R1	R2	R3	R4	R5	S3	S4	250	250
E	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	500	500
F	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	1000	1000
G	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	2000	2000
H	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	4000	4000

Key:
S1 – S14 and R0 – R5 are as defined for Template 1 (see above).
Values in columns 11 and 12 are concentrations in pg/ml for IL-6.

7.10.3 Addition of substrate solution and measuring

Prepare the substrate solution shortly before use. Transfer 90 ml of substrate buffer (Reagent 6) to a plastic bottle, add 4.5 ml of TMB solution (Reagent 5), and mix.
Wash the microtitre plates by immersing them 3 times in wash solution, then rinse 3 times with demineralised water. Place the microtitre plates face down on cellulose and tap.
Pipette 200 µl of substrate solution into each well. After 10 - 15 minutes, stop the enzyme reaction by adding 50 µl/well of stopping solution. Wipe the back of the microtitre plates with a clean tissue, then measure the absorbance at 450 nm in a ELISA plate reader using a 540 nm to 590 nm corrective filter. Subtract the values of the measurement with the corrective filter from values measured with the 450 nm filter.

8 DATA ANALYSIS, PREDICTION MODEL AND RELATED ERRORS

8.1 Acceptance Criteria

The assay should be considered acceptable only, if the following criteria are met:

The ELISA is valid if the OD of the blank control is below 0.15 and the mathematical function (quadratic model) of the IL-6 standard curve produces an r >0.95.

The reactions (in terms of OD) on the endotoxin concentrations give a sigmoidal ascending dose response.

Blood donors are considered low responders if their mean OD value for the endotoxin reference standard concentration 1 EU/ml (R5) is below the mean OD value for 1000 pg/ml of IL-6. Low responders should not be included in the assessment.
Blood donors showing a mean OD value for the negative control (R0) above the mean OD value at 500 pg/ml of IL-6 per millilitre are also not included in the assessment (high responder).
If the test sample show an irregular response (e.g. high SD), check the results obtained for the blank and the standard endotoxin concentrations derived from the donor in question. If the latter results are inconsistent with those which would ordinarily be expected, then the donor in question must be excluded from the assessment.

8.2 Interference Test

Interference with the cell system

For each product tested the first time, it is necessary to determine whether it requires dilution prior to assay or not. The following experiment checks for interference between the sample and the PBMC and/or ELISA system.
Perform a dilution series of the product with dilution factor 2 or higher, if necessary. Dilutions should not reach outside the range where the detection limit of the test does not allow to determine the defined **endotoxin limit** of the product. If no endotoxin limit is defined it can easily be estimated by dividing 350EU by the maximum hourly dose. (Example: The maximum hourly dose is 100mg/patient; then the estimated endotoxin limit is 350/100 = 3.5EU/mg)
Split each dilution ("undiluted" may be included) and test it spiked and unspiked in quadruplicates. Spike with an endotoxin concentration from the middle of the endotoxin standard curve (e.g. 0.25 EU/ml). Prepare the spiking solution in RPMI-C instead of saline and prefill the wells for this purpose first with 50µl of RPMI-C and then add 50µl of the spiking solution. The prefill for the unspiked testing of the dilution is 100 µl of RPMI-C per well. Then add 50µl of the product dilution (or undiluted product, resp.) to 4 wells containing the spiking solution and 4 wells containing only RPMI-C.

Proceed according to the testing instructions (see above).
Determine the lowest dilution (highest concentration) of the product that yields an endotoxin spike recovery of 50 to 200%. For that purpose calculate the mean endotoxin values of the unspiked and the spiked product dilutions applying a best fit model on the endotoxin calibration curve. Consider that in this nomenclature "undiluted" means also a dilution. Then subtract the endotoxin value of the unspiked dilution from the endotoxin value of the corresponding spiked dilution. Calculate the spike recovery for each dilution in percent taking the theoretical value (spike concentration e.g. 0.25 EU/ml) as a 100%.
Example: Tested dilutions and spike recovery :"undiluted" = 25%; 1 in 2 = 49%; 1 in 4 = 90% and 1 in 8 = 110%. Then the dilution meeting the requirements is 1 in 4.

8.3 Interference with the ELISA system

Simulate a pyrogen test without cells and without incubation using the chosen test dilution of the product (from testing "interference with the cell system", in the example 1 in 4). Prepare enough replicates to spike with the IL-6 standard curve concentrations (4000 - 62.5 pg/ml and the blank) in duplicate (12 wells). Test the IL-6-spiked solutions directly for interference in the ELISA system together with a unspiked IL-6 standard dilution series. Apply exactly the same methodology as for a real pyrogen test. If interference in the ELISA system appears, the product has to be tested in further dilutions but not exceeding the maximum valid dilution (see above).
The lowest dilution (highest concentration) of the product not interfering with PBMC and the ELISA is considered as the future dilution for routine testing.

8.4 Prediction model

Perform the test with a certain product according to this SOP. Apply a valid test dilution showing no interference with the test systems as written above. Calculate the parameters of the endotoxin standard curve applying a best fit model ($r \geq 0.95$). Reject

outliers only after checking according to Dixon's Test. Test at the level of p = 0.90 or 0.95. Measure the mean OD values of all replicates (at least triplicates). Using the endotoxin standard curve, calculate the pyrogen content of the product in endotoxin equivalents. Multiply the calculated concentration of endotoxin equivalents in the product by the dilution factor (may be 1 in special cases). This value represents the pyrogen content of the sample expressed in endotoxin equivalents for the donor under test. For the product to pass the test, it must comply with its specification (endotoxin limit concentration) when tested with PBMC from three independent donors.

Remark: In order to optimize the prediction model of this test a methodology using a well characterized reference substance (identical in composition to the test substance) is in preparation. With this advanced prediction model interferences in the test due to donor variability will be reduced distinctively.

9 HEALTH SAFETY AND ENVIRONMENT

Human material

Human material should be treated as biologically hazardous and all work using human material is to be carried out according to the procedures specified in the NOVARTIS Safety Guidelines.

Cultures of human material should be treated as biologically hazardous waste and disposed of according to the procedures specified in the NOVARTIS Safety Guidelines.

Bacterial endotoxin is, as its name indicates, a toxic agent and should be handled with care.

Precautions: Cover open cuts before use. Do not get in eyes, on skin, on clothing. Avoid inhaling. Keep container closed.

First Aid: In case of contact, immediately flush eyes or skin with plenty of water for at least 15 minutes. If inhaled, remove to fresh air. If not breathing, give artificial respiration, preferably mouth-to-mouth. If breathing is difficult, give oxygen.

Effects of skin absorption can include fever, headache and hypotension.

Effects of inhalation can include fever, headache and hypotension.

Effects of ingestion - adverse effects are unlikely since ingested endotoxin is rapidly detoxified.

10 REFERENCES

Bleeker, W.K., de Groot, E.M., den Boer, P.J., Biessels, P.T., Aarden, L.A. and Bakker, J.C. (1994) Measurement of interleukin-6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol. 22, 835-40.

Dinarello, C.A., Gatti, S. and Bartfai, T. (1999) Fever: links with an ancient receptor. Curr Biol. 9, R147-50.

Dinarello, C.A., O'Connor, J.V., LoPreste, G. and Swift, R.L. (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J. Clin. Microbiol. 20, 323-329.

Duff, G.W. and Atkins, E. (1982) The detection of endotoxin by *in vitro* production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52, 323-332.

Fennrich, S., Fischer, M., Hartung, T., Lexa, P., Montag-Lessing, T., Sonntag, H.G., Weigandt, M. and Wendel. (1999) A Detection of endotoxins and other pyrogens using human whole blood. Dev Biol Stand. 101, 131-9.

Finney DJ: Statistical Method in Biological Assay, Third Edition. London,
Charles Griffin and Company Ltd, 1978.

Gaines Das RE and Tydeman MS. (1982) Iterative weighted regression analysis of
logit responses: A computer program for analysis of bioassays and
immunoassays. Computer Programs in Biomedicine 15; 13-22.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15, 425-33.

Levin, J. and Bang, F.B. (1964) A description of cellular coagulation in the Limulus. Bull. John Hopkins Hosp. 155, 337-345.

Mascoli, C.C. and Weary, M.E. (1979a) Limulus amebocyte lysate (LAL) test for detecting pyrogens in parenteral injectable products and medical devices: advantages to manufacturers and regulatory officials. J Parenter Drug Assoc. 33, 81-95.

Mascoli, C.C. and Weary, M.E. (1979b) Applications and advantages of the Limulus amebocyte lysate (LAL) pyrogen test for parenteral injectable products. Prog Clin Biol Res. 29, 387-402.

Meager, A., Parti, S., Leung, H., Peil, E. and Mahon, B. (1987) Preparation and characterization of monoclonal antibodies directed against antigenic determinants of recombinant human tumour necrosis factor (rTNF). Hybridoma 6, 305-311.

Poole, S., Selkirk, S., Rafferty, B., Meager, A., Thorpe, R. and Gearing, A. (1989) Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa, November 1989, 17-18.

Poole, S., Thorpe, R., Meager, A., Hubbard, A.R. and Gearing, A.J.H. (1988) Detection of pyrogen by cytokine release. Lancet 8577, 130.

Ray, A., Redhead, K., Selkirk, S. and Poole, S. (1990) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211-218.

Schins, R.P.F., van Hartingsveldt, B. and Borm P.J. (1996) *Ex vivo* cytokine release from whole blood. Exp. Toxic Pathol. 48, 494-496.

Taktak, Y.S., Selkirk, S., Bristow, A.F., Carpenter, A., Ball, C., Rafferty, B. and Poole, S. (1991) Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43, 578-582.

Validation of Biomedical Testing Methods

In Vitro Pyrogen Test Using PBMC

(SP+PB var. Novartis) 03 10 02

Standard Operating Procedure

Only the responsible of the GLP/QA Unit is allowed to make copies of this document. Extra examples can be obtained at the GLP/QA Unit. Quality Documents are only valid if they are signed by the responsible of the GLP/QA Unit in and provided with a blue copy number	Copy number

Sop-PBMCv02

JOINT RESEARCH CENTRE
EUROPEAN COMMISSION
Institute for Health and Consumer Protection
ECVAM

In Vitro Pyrogen Test Using Peripheral Blood Mononuclear Cells

Version number: v02
Applicable from: _____
Expired at: _____

Drafted by: Name Peter Bruegger
Date
Signature

Reviewed by: Name
Date
Signature

Approved by: Name
Date
Signature

Issued by: Name
Date
Signature

*Owner/Trainer: Signature: Date:

Institute for Health and Consumer Protection

PAGE OF CHANGES

Date of change/ Date of draft:	Version-number:	Changed page(s):	Summary of the change(s):	Changed by/Sign.:
17. Sep. 2001				
11. April 2002			PBMNC to PBMC Endotoxin standard conc. Chapter 8 Annex 1 deleted	
3. Oct. 2002			Sample dilution to MVD Extension to 4 donors per test Extension of prediction model	

TABLE OF CONTENTS ..... Page No.

1 INTRODUCTION 5
2 PURPOSE 5
3 SCOPE / LIMITATIONS 5
4 METHOD OUTLINE 6
5 DEFINITIONS / ABBREVIATIONS 7
6 MATERIALS 8
6.1 Buffers and Reagents for Novartis IL-6 ELISA: 9
7 METHODS 12
7.1.1 Coating of IL-6 ELISA plates 12
7.1.2 Preparation of samples for assay [at] 12
7.1.3 Collection of human blood 13
7.1.4 Isolation of PBMCs [at] 13
7.1.5 Washing PBMCs [at] 13
7.1.6 Equilibration of reagents for cell culture 14
7.1.7 Preparation of the LPS standard curve [at] 14
7.2 Cell Culture [at] 15
7.3 Detection of IL-6 in the supernatant medium by ELISA 16
7.3.1 Preparation of IL-6 standard curve 17
7.3.2 Addition of standards and samples 17
7.3.3 Addition of substrate solution and measuring 18
8 DATA ANALYSIS, PREDICTION MODEL AND RELATED ERRORS 19
8.1 Acceptance Criteria 19
8.2 Interference Test 19
8.2.1 Interference with the cell system 19
8.2.2 Interference with the ELISA system 20
8.3 Prediction model 20
9 HEALTH SAFETY AND ENVIRONMENT 22
10 ANNEX 23
11 REFERENCES 24

1 INTRODUCTION

Parenteral pharmaceutical products must be shown to be free from pyrogenic (fever-inducing) contamination. While a pyrogen may in general be defined as any substance that causes fever, the pyrogens that almost invariably contaminate parenteral pharmaceuticals are bacterial endotoxins (lipopolysaccharides, LPS) from Gram-negative bacteria (Mascoli and Weary, 1979a, 1979b). There are two Pharmacopoeial tests for pyrogenic contamination: the rabbit pyrogen test and the Limulus amoebocyte lysate (LAL) test. The rabbit pyrogen test, which detects LPS and other pyrogens, involves measuring the rise in body temperature evoked in rabbits by the intravenous injection of a sterile solution of the substance to be examined. In contrast, the LAL test detects only LPS: it is described in Pharmacopoeias as the bacterial endotoxins test (BET). The principle of the LAL-test is that LPS causes extracellular coagulation of the blood (haemolymph) of the horseshoe crab, *Limulus polyphemus.* (Levin & Bang, 1964). Although the LAL test is gradually superseding the rabbit pyrogen test, hundreds of thousands of rabbit pyrogen tests are still carried out each year around the world, largely on products which cannot, for one reason or another, be tested in the LAL test. While proving generally reliable, both the rabbit pyrogen test and LAL test have shortcomings. The rabbit pyrogen test uses experimental animals, is costly and is not quantitative. The LAL test gives false negatives with certain products, can overestimate the pyrogen content of other products and does not detect pyrogens other than bacterial endotoxin (LPS), such as Gram-positive exotoxins, viruses and fungi (Dinarello et al., 1984; Poole et al., 1988; Ray et al., 1990; Taktak et al., 1991; Fennrich et al., 1999).
The basis of the rabbit pyrogen test is the in vivo stimulation by exogenous pyrogens (usually LPS) of rabbit peripheral blood monocytes to produce the endogenous pyrogens that cause fever. The endogenous pyrogens are pyrogenic cytokines such as tumour necrosis factorα (TNFα), interleukin-1 (IL-1α and IL-1β, two separate gene products), IL-6 and IL-8 (Dinarello et al., 1999). In view of the shortcomings of the rabbit pyrogen test and the LAL test, in vitro pyrogen tests that utilise the exquisite sensitivity to exogenous pyrogen of monocytes have been proposed. In such tests, products are incubated with human peripheral blood monocytes (or mononuclear cells, PBMC, or leukocytes) and the conditioned media assayed for pyrogenic cytokines (Duff & Atkins, 1982; Dinarello et al., 1984; Poole et al., 1988; Poole, 1989; Hansen and Christensen, 1990; Taktak et al., 1991; Bleeker et al., 1994).

2 PURPOSE

By quantifying the amounts of cytokines released by PBMC stimulated with the USP reference preparation for endotoxin or the international standard (IS) for endotoxin (LPS), various non-endotoxin pyrogens and with medicinal products spiked with endotoxin, it is the objective that an in vitro pyrogen test be developed that will serve as a replacement for the rabbit pyrogen test.

3 SCOPE / LIMITATIONS

The method described below is for the evaluation of in vitro PBMC/cytokine release tests using IL-6 as the cytokine that serves as the readout (measured variable). It is not a 'finalised' test system for the testing of medicinal products. Especially the prediction model has to be developed cautiously.

The method may be applied only to preparations that have been validated with the method, i.e. shown not to interfere by causing inhibition or enhancement of LPS (STD)-induced cytokine production.

4 METHOD OUTLINE

Freshly taken human whole blood is heparinised, diluted with phosphate buffered saline (PBS) and the PBMC's isolated and stimulated for 16-24h with the USP reference preparation for endotoxin, the international standard for endotoxin, (LPS) or samples of related materials, e.g. other endotoxins, non-endotoxin pyrogens and medicinal products unspiked and spiked with endotoxin. Following this stimulation, the concentrations of the cytokine in the PBMC-conditioned medium are quantified using a specific ELISA for IL-6 (which is calibrated in terms of the appropriate international standard). The construction of dose-response curves for endotoxin (LPS) versus optical densities of released cytokin permits the estimation of the pyrogenic/cytokine-releasing activity contained in the samples.

5 DEFINITIONS / ABBREVIATIONS

Ab	Antibody
BSA	Bovine serum albumin
CO_2	Carbon dioxide
°C	Degrees Celsius (Centigrade)
D-R	Dose-response
E. coli	Escherichia coli
ELISA	Enzyme-linked immunosorbent assay
EP	European Pharmacopoeia
EU	Endotoxin units
FDA	Food and Drug Administration (USA)
g	Gram
h	Hour
HIFCS	Heat-inactivated (+56°C for 30 min) foetal calf serum
H_2O_2	Hydrogen peroxide
H_2SO_4	Sulphuric acid
IL	Interleukin
IS	International standard
IU	International unit
l	Litre
KOH	Potassium hydroxide
LAL	Limulus amoebocyte lysate
LPS	Lipopolysaccharide
M	Molar
MAb	Monoclonal antibody
mg	Milligram
min	Minute
ml	Millilitre
mM	Millimolar
µg	Microgram
µl	Microlitre
NaCl	Sodium chloride
NaOH	Sodium hydroxide
$NaHCO_3$	Sodium hydrogen carbonate
NaH_2PO_4	Sodium di-hydrogen phosphate
Na_2HPO_4	di-Sodium hydrogen orthophosphate
nm	Nanometre
OD	Optical density
PBMC	Peripheral blood mononuclear cells
PBS	Dulbecco's Phosphate Buffered Saline
PF	Pyrogen-free (items purchased as sterile and pyrogen-free or baked at 250° for 30-60 min.
POD	Horseradish peroxidase conjugate
R	Endotoxin standard
rpm	Rounds per minute
RPMI	RPMI 1640 cell culture medium
RPMI-C	RPMI 1640 cell culture medium + human AB serum at a final concentration of 2% v/v
RSE	Reference Standard Endotoxin
RT	Room temperature

TMB	Tetramethyl benzidine
S	Test sample
STD	Standard
UNK	Unknown
USP	United States Pharmacopoeia
x g	x gravity

6 MATERIALS

Sterile, pyrogen-free Dulbecco's phosphate buffered saline w/o Ca^{++} and Mg^{++} (Life Technologies)
Polyoxyethylene-sorbitan monolaurate (TWEEN 20), cell culture grade, (Sigma, P-2287)
Hydrochloric acid, 0.1M, sterile filtered (Sigma, H-9892)
Sodium hydroxide (reagent grade)
1M H_2SO_4 (Merck)
Mouse monoclonal anti-IL-6 antibody from clone 16
Horseradish peroxydase conjugated sheep polyclonal anti-IL-6 antibody
3,3',5,5'-Tetramethyl benzidine (e.g. Fluka Cat. No. 87748)
Acetone (reagent grade)
Ethanol (reagent grade)
Phenole (e.g. Merck Cat. No. 100206)
Potassium hydroxide (reagent grade)
Sodium dihydrogen phosphate (e.g. Merck Cat. No. 106346)
Disodium hydrogen phosphate (e.g. Merck Cat. No. 106580)
Tris (hydroxymethyl) aminomethane (e.g. Fluka Cat. No. 93352)
Kathon MW/WT, Christ Chemie AG, Reinach, Switzerland
Albumin from bovine serum (e.g. Fluka Cat. No. 05480)
Citric acid monohydrate e.g. Fluka Cat. No. 27490)
Human AB serum (Sigma)
Trypan blue stain (Sigma)
USP Reference Standard Endotoxin [EC6 lot G], identical to the WHO international standard for bacterial endotoxin (LPS, vial code 94/580)
Fragmin (Dalteparin, 10000 IU/ml, Pharmacia)
RPMI 1640 medium (Life Technologies™, Paisley, Scotland)
L-Glutamine 200mM (Life Technologies™, Paisley, Scotland)
Penicillin/Streptomycin solution (Seromed Cat. No. A2213)
Lymphoprep (Nycomed, Oslo, Norway)
Nunc-Immuno 96-well plate MaxiSorp (F96, Life Technologies™, Paisley, Scotland)
Falcon Microtest tissue culture plate, 96-well (353072, Beckton Dickinson Labware)
Falcon serological pipettes (5ml,10ml, 25ml, Beckton Dickinson Labware)
Centrifuge tubes (Falcon 2070 Blue Max™)
Polypropylene conical tubes (Falcon 2069 Blue Max™)
Eppendorf Biopur Tips 100ul & 1000ul (Eppendorf-Netheler-Hinz-Gmbh,Germany)
0.22 μm sterile filters (MilliPak 60, Millipore)
Eppendorf® volumetric pipettes

All other consumables are purchased as sterile and pyrogen-free and other reagents are pro analysis grade.

6.1 Buffers and Reagents for Novartis IL-6 ELISA:

Coating Buffer

Dissolve
5.0 g of sodium dihydrogen phosphate, (e.g. Merck Art. No. 106346), and
2.9 g of disodium hydrogen phosphate, (e.g.Merck Art. No. 106580),
in 400 ml of distilled water.

Use 1 N NaOH to adjust the pH to 7.5, and make up to 500 ml with distilled water.

Remains stable for 6 months at 2 - 8 °C.

Blocking Buffer

Tris(hydroxymethyl)aminomethane, (e.g.Fluka Art. No. 93352)	12.1 g
Dissolve in distilled water	400 ml
Kathon MW/WT, Christ Chemie AG, Switzerland	0.1 ml

Use 4 M HCl to adjust the pH to 7.5.

Albumin from bovine serum, (e.g.Fluka Art. No. 05480)	5.0 g
Add distilled water to make up to 500 ml.	

Remains stable for 6 months at 2 - 8 °C.

Stopping Solution

Distilled water	500 ml
H_2SO_4	26.6 ml

Wash Solution

Demineralised water	2000 ml
Tween-20	1 ml

Interleukin-6 Standard

Use IL-6 from human lymphocytes (e.g. Boehringer Mannheim, Cat. No. 1299972). 1 vial contains 200 000 units of natural human IL-6.
Dilute the contents of a vial (1 ml) of human IL-6 with 4 ml of RPMI-C, and freeze aliquots of 100 µl at about – 80 °C.
Before the first use of a new batch, this IL-6 used as the assay calibrant has to be calibrated against the IS for IL-6 (89/548).

Mouse Anti-Human IL-6 Monoclonal Antibody

(*Reagent 1)*
derived from clone 16.

Detection Antibody (POD)

(Reagent 2)
Sheep anti-human IL-6 antibody, horseradish peroxidase conjugated. Remains stable for at least 6 months at 2 - 8 °C.

Human IL-6 Standard Solution

(Reagent 3)
Thaw 1 frozen aliquot of interleukin-6 standard (4000 units in 100 µl) and dilute with 900 µl of RPMI-C.
Add 100 µl of this dilution to 900 µl of RPMI-C = 400 units/ml (standard initial concentration, equivalent to 4000 pg/ml).
Prepare the solution shortly before use; do not store.

Dilution Buffer

(*Reagent 4)*
Prepare the dilution buffer as follows:

Tris(hydroxymethyl)aminomethane, (e.g.Fluka Art. No. 93352)	2.1 g
Distilled water	400 ml
Kathon MW/WT, Christ AG, Switzerland	0.1 ml
Phenol, (e.g.Merck Art. No. 100206)	0.5 g
Heat-inactivated (30 minutes/56 °C) fetal bovine serum	25 ml

Mix to dissolve the substances, then adjust the pH to 7.5 with 4 M HCl. Make up to 500 ml with distilled water.

Remains stable for at least 6 months at 2 - 8 °C.
In the absence of the stabilizers Kathon and phenol the stability is only 1 day.

TMB Solution

(Reagent 5)

Prepare the TMB solution as follows:

3,3',5,5'Tetramethylbenzidine (e.g.Fluka Art. No. 87748)	240 mg
Reagent-grade acetone	5 ml

Dissolve, then add

Reagent-grade ethanol	45 ml
Perhydrol (30 % H_2O_2), (e.g.Merck Art. No. 107209)	0.300 ml

Remains stable for at least 6 months at 15 - 25 °C when sealed and protected from light.

Substrate Buffer

(Reagent 6)

Reagent-grade citric acid monohydrate (e.g.Fluka Art. No. 27490)	6.3 g
Distilled water	800 ml

Mix to dissolve, then adjust the pH to 4.1 by adding 4 M KOH.
Make up to 1000 ml with distilled water and add 0.2 ml of Kathon MW/WT

Remains stable for about 6 months at 15 - 25 °C.
In the absence of the Kathon the stability is only 1 day.

Culture Medium (RPMI-C)

The preparation of the medium has to be carried out in a class 2 laminar flow sterile cabinet, using aseptic technique and reagents and consumables that are sterile and pyrogen-free.

RPMI 1640 medium	500 ml
Human serum AB	5 ml
L-Glutamine, 200 mM	5 ml
Penicillin/Streptomycin solution	10 ml

7 METHODS

Steps marked (^{at}) are carried out in a class 2 laminar flow sterile cabinet, using aseptic technique and reagents and consumables that are sterile and pyrogen-free.

7.1.1 Coating of IL-6 ELISA plates

Dilute the coating anti-IL-6 antibody (Clone 16) with coating buffer to 2.5 µg/ml and swirl to mix, e.g. 1 mg of antibody in 400 ml of coating buffer. Add 200 µl to each well of a 96-well plate (Nunc-Immuno MaxiSorp F96)
Stack the microtitre plates and allow to stand in the dark at 15 - 25 °C.for 16 - 24 hours.

Aspirate and discard the coating solution. Wash the coated plate 3 times with demineralised water, and tap out onto cellulose.
Pipette 200 µl of blocking buffer into each of the wells to block the residual protein-binding capacity of the coated plates.

Seal the plates with adhesive film, and store in a humidified atmosphere at 2 - 8 °C (shelf life: two months).

7.1.2 Preparation of samples for assay ^{at}

Samples are tested at the MVD according to the "Trial Plan".

7.1.3 Collection of human blood

Qualification of blood donors: Blood donors are to describe themselves as being in good health, not suffering from any bacterial or viral infections (including colds and influenza). Blood donors are not to be taking drugs known to influence the production of cytokines. Also, See Section 8. (below) for the criteria for the rejection of data as having come from a test utilising blood from a donor not in good health.

***Procedure*[at]:** Dilute 1 ampoule of Fragmin (10000 IU in 1 ml) with 1 ml pyrogen-free distilled water. Draw back the piston of a 30 ml syringe by about 3 mm, then inject, using a syringe calibrated in µl, 50 µl of the diluted Fragmin via the luer lock adapter into the space between the syringe wall and the piston. Then attach a 19 mm, 21 gauge butterfly system to the 30 ml syringe. Alternatively a 40mm 21 gauge needle may be used.

Using this prepared (heparinised) syringe, draw 30 ml blood from the median cubital or cephalic vein of the left or right arm of a single donor. Mix the blood with the anticoagulant by tilting the syringe several times.

Afterwards remove the butterfly or the needle resp. and transfer 15 ml of the blood into each of two 50 ml sterile, pyrogen-free centrifuge tubes.

Each single in vitro pyrogen test has to be executed with the cells of four donors individually.

7.1.4 Isolation of PBMCs [at]

PBMCs are isolated using Lymphoprep. The procedure is a modification of the manufacturers' instructions that may permit cleaner separation of the PBMC. Start the procedure not later than 2 hours after blood withdrawal.

Procedure: Add 15 ml PBS to each 15 ml of heparinised whole blood to make 2 tubes of 30 ml of diluted blood. Remove the cotton from a 10-ml pipette and place the pipette carefully (point down) in the tube with the diluted blood. Add 20 ml Lymphoprep to each of the two tubes, via the inserted pipette to form a lower layer. Centrifuge at 340 x g for 45 min at room temperature (18 – 25 °C), with the centrifuge brake set to off/zero. After centrifuging, the monocytes form a white band at about the 25 ml graduation of the tube. Carefully draw off the supernatant phase and discard. Take up the PBMC with a 10 ml pipette and transfer to a new 50 ml centrifuge tube.

7.1.5 Washing PBMCs[at]

To the isolated PMBNC add sufficient PBS to give a total volume of 50 ml and centrifuge at 340 x g for 15 min.

Pour off the supernatant phase, and resuspend the sediment with 10 ml of PBS using a serological pipette (aspirate and expell several times, do not vortex). Make up to a total volume of 50 ml with PBS and centrifuge at 340 x g for 10 min.

Pour off the supernatant phase and resuspend the sediment with 15 ml of RPMI-C. Pool the resuspended sediments of both tubes in one tube. Determine the cell count per ml in a Neubauer haemocytometer (see literature haematology). Dilute the cell suspension with RPMI-C to 1 mio cells/ml if the cell count is above 1.2 mio cells/ml. This suspension of PBMC in RPMI-C is used in the cell-culture .

The cells shall be stimulized by endotoxin or samples within 4 hours after blood withdrawal.

7.1.6 Equilibration of reagents for cell culture

Bring a vial of the LPS standard, the samples for assay and a bottle of RPMI-C to room temperature.

7.1.7 Preparation of the LPS standard curve [at]

Add 5 ml of LAL reagent water to the lyophilised contents of one vial of the current USP Reference Standard Endotoxin or the IS to produce a stock solution of 2000 EU(IU)/ml. Vortex for at least 30 min. The stock solution remains stable for 14 days if stored at 2 – 8°C.
Prepare the LPS standard curve by making serial dilutions in saline of the stock solution of endotoxin:

Label seven tubes, A – G. Add the volumes of saline to the tubes specified in table 1, below.

Take 200µl of endotoxin stock solution and add 1.8 ml of saline and vortex to make 2.0 ml of a 200 EU (IU)/ml solution of LPS = Solution S.

Table 1. Preparation of the LPS standard curve:

Tube	LPS added to tube	Saline	[LPS] in tube	→ *[LPS] in well*
A	100 µl of Solution S = 20 EU	900 µl	20 EU/ml	*Not for culture*
B	100 µl of Solution A = 2 EU	1900 µl	1 EU/ml	*0.2 EU/ml*
C	500 µl of Solution B = 0.5 EU	500 µl	0.5 EU/ml	*0.1 EU/ml*
D	500 µl of Solution C = 0.25 EU	500 µl	0.25 EU/ml	*0.05 EU/ml*
E	500 µl of Solution D = 0.125 EU	500 µl	0.125 EU/ml	*0.025 EU/ml*
F	500 µl of Solution E = 0.063 EU	500 µl	0.063 EU/ml	*0.0125 EU/ml*
G	None	1 ml	0 EU/ml	*0 EU/ml*

Vortex each of Solutions A - G after its preparation and then use each solution for the preparation of the subsequent dilution.
(Multiples of the above volumes may be used to generate larger volumes of Solutions A – G and an LPS standard calibrated against theUSP RSE or the IS for endotoxin may be used.)

7.2 Cell Culture [at]

Add 100 µl of RPMI-C to wells of columns 1 – 10 as in template 1, see below.

Add 50 µl of the test samples S1 – S14 to wells as in template 1, see below.

Add 50 µl of LPS standards to wells as in template 1, below.

Solution G into wells A3 – D3 (STD R0)
Solution F into wells A4 – D4 (STD R1)
Solution E into wells A5 – D5 (STD R2)
Solution D into wells A6 – D6 (STD R3)
Solution C into wells A7 – D7 (STD R4)
Solution B into wells A8 – D8 (STD R5)

(The above order of addition permits the same tip to be used for additions of all the standards.)

Gently swirl the solution of PBMC to reduce settling of the cells and to distribute the PBMC more evenly throughout the RPMI-C solution immediately before aliquots of PBMC are taken. Do not vortex.

Add 100 µl of PBMC to the 96-well plate (see template 1, below). Add the PBMC by row in the following sequence: A, E, B, F, C, G, D, H. A repeating pipette may be used for these additions provided that the aliquots are added briskly to minimise the settling of cells.

Gently swirl the resulting cultures to mix the contents of the wells without cross-contaminating wells.
Incubate the cultures without vibration (to allow the cells to settle) at 37°C for 16 - 24h in an atmosphere of 5% CO_2 in humidified air.

Template 1: PBMC culture plate

	1	2	3	4	5	6	7	8	9	10	11	12
A	S1	S2	R0 0	R1 0.063	R2 0.125	R3 0.25	R4 0.5	R5 1	S3	S4	void	void
B	S1	S2	R0 0	R1 0.063	R2 0.125	R3 0.25	R4 0.5	R5 1	S3	S4	void	void
C	S1	S2	R0 0	R1 0.063	R2 0.125	R3 0.25	R4 0.5	R5 1	S3	S4	void	void
D	S1	S2	R0 0	R1 0.063	R2 0.125	R3 0.25	R4 0.5	R5 1	S3	S4	void	void
E	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	void	void
F	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	void	void
G	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	void	void
H	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	void	void

Key:
S1 – S14 = test samples #1 - #14
R0 – R5 = Reference standard for endotoxin , R0 = 0 EU/ml, R1 = 0.063 EU/ml, R2 = 0.125 EU/ml, R3 = 0.25 EU/ml, R4 = 0.5 EU/ml and R5 = 1 EU/ml (The final concentrations are: 0.0125, 0.025, 0.05, 0,1 and 0,2UE/ml).

7.3 Detection of IL-6 in the supernatant medium by ELISA

Immunoreactive IL-6 in aliquots of the tissue culture fluid is quantified using a validated ELISA, in which the IL-6 standard used as the assay calibrant is calibrated against the IS for IL-6 (89/548).

7.3.1 Preparation of IL-6 standard curve

Prepare the IL-6 standard dilutions as follows:
Fill each of 7 polystyrene tubes (12ml) with 500 µl of RPMI-C.
Add 900 µl of RPMI-C to one frozen aliquot (100µl) of IL-6 standard (=Reagent 3)
Dilute this concentration (Solution I) by transferring 500 µl to tube J (1 in 2). Further dilute by transferring 500 µl from this tube to the next, 500 µl from that tube to the following one, and so on, ending with the tube marked with O. Use the RPMI-C in tube P as a blank (see table 2).

Table 2. Preparation of the IL-6 standard curve from an aliquot of the IL-6 standard.

Sol.	IL-6 added	RPMI-C	IL-6 in tube
I	100 µl of frozen sol.= 4000 pg	900 µl	4000 pg/ml
J	500 µl of Solution I = 2000 pg	500 µl	2000 pg/ml
K	500 µl of Solution J = 1000 pg	500 µl	1000 pg/ml
L	500 µl of Solution K = 500 pg	500 µl	500 pg/ml
M	500 µl of Solution L = 250 pg	500 µl	250 pg/ml
N	500 µl of Solution M = 125 pg	500 µl	125 pg/ml
O	500 µl of Solution N = 62.5 pg	500 µl	62.5 pg/ml
P	None	500 µl	0 pg/ml

Vortex each of Solutions I – O after its preparation and then use each solution for the preparation of the subsequent dilution.

(*Multiples of the above volumes may be used to generate larger volumes of Solutions I – P).*

7.3.2 Addition of standards and samples

Volumes are sufficient for 4 microtiter plates.
Dilute detection antibody (Reagent 2) with dilution buffer (Reagent 4) (e.g. 1 in 500 [200 µl of Reagent 2 + 100 ml of Reagent 4]), and mix without causing foam to form Test each batch of detection antibody in separate experiments to determine the optimum dilution.

Just before using shake out the blocking buffer from the antibody-coated microtitre plates, place the plates on cellulose with the openings facing down, and tap.
At the end of the tissue culture incubation, transfer 50 µl of supernatant from each of the wells of columns 1 – 10 of the tissue culture plate to the corresponding wells on each of the cytokine ELISA plates – see template 1, above and template 2, below. A multichannel pipette may be used. Ensure that the well contents are mixed by aspirating and expelling 50 µl three times before transferring the liquid. (The wells in columns 11 and 12 are for the IL-6 standard curve – see, below).

Transfer 50 µl of each of the dilutions of the IL-6 standard, and of the blank, into 2 wells each (standard concentrations from 4000 - 62.5 pg/ml).
Add 200 µl of dilute detection antibody (e.g. 1 in 500) to each of the wells, seal the microtitre plates with adhesive film, and allow to stand for 2 - 3 hours at 20 - 25 °C. Use template 2 below.

Template 2: ELISA plate

	1	2	3	4	5	6	7	8	9	10	11	12
A	S1	S2	R0	R1	R2	R3	R4	R5	S3	S4	0	0
B	S1	S2	R0	R1	R2	R3	R4	R5	S3	S4	62.5	62.5
C	S1	S2	R0	R1	R2	R3	R4	R5	S3	S4	125	125
D	S1	S2	R0	R1	R2	R3	R4	R5	S3	S4	250	250
E	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	500	500
F	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	1000	1000
G	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	2000	2000
H	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	4000	4000

Key:
S1 – S14 and R0 – R5 are as defined for Template 1 (see above).
Values in columns 11 and 12 are concentrations in pg/ml for IL-6.

7.3.3 Addition of substrate solution and measuring

Prepare the substrate solution shortly before use. Transfer 90 ml of substrate buffer (Reagent 6) to a plastic bottle, add 4.5 ml of TMB solution (Reagent 5), and mix.
Wash the microtitre plates by immersing them 3 times in wash solution, then rinse 3 times with demineralised water. Place the microtitre plates face down on cellulose and tap.
Pipette 200 µl of substrate solution into each well. After 10 - 15 minutes, stop the enzyme reaction by adding 50 µl/well of stopping solution. Wipe the back of the microtitre plates with a clean tissue, then measure the absorbance at 450 nm in a ELISA plate reader using a 540 nm to 590 nm corrective filter. Subtract the values of the measurement with the corrective filter from values measured with the 450 nm filter.

8 DATA ANALYSIS, PREDICTION MODEL AND RELATED ERRORS

8.1 Acceptance Criteria

The assay should be considered acceptable only, if the following criteria are met:

The ELISA is valid if the OD of the blank control is below 0.15 and the mathematical function (quadratic model) of the IL-6 standard curve produces an r^2 >0.95.

The reactions (in terms of OD) on the endotoxin concentrations give a sigmoidal ascending dose response.

Blood donors are considered low responders if their mean OD value for the endotoxin reference standard concentration 1 EU/ml (R5) is below the mean OD value for 1000 pg/ml of IL-6. Low responders should not be included in the assessment.
Blood donors showing a mean OD value for the negative control (R0) above the mean OD value at 500 pg/ml of IL-6 per millilitre are also not included in the assessment (high responder)
If the test sample show an irregular response (e.g. high SD), check the results obtained for the blank and the standard endotoxin concentrations derived from the donor in question. If the latter results are inconsistent with those which would ordinarily be expected, then the donor in question must be excluded from the assessment.

From a set of four donors a maximum of one donor may be excluded from the assessment, otherwise the test has to be repeated with four different donors.

8.2 Interference Test

8.2.1 Interference with the cell system

For each product tested the first time, it is necessary to determine whether it requires dilution prior to assay or not. The following experiment checks for interference between the sample and the PBMC and/or ELISA system.
Perform a dilution series of the product with dilution factor 2 or higher, if necessary. Dilutions should not reach outside the range where the detection limit of the test does not allow to determine the defined **endotoxin limit** of the product (MVD). If no endotoxin limit is defined it can easily be estimated by dividing 350EU by the maximum hourly dose. (Example: The maximum hourly dose is 100mg/patient; then the estimated endotoxin limit is 350/100 = 3.5EU/mg)
Split each dilution ("undiluted" may be included) and test it spiked and unspiked in quadruplicates. Spike with an endotoxin concentration from the middle of the endotoxin standard curve (e.g. 0.25 EU/ml). Prepare the spiking solution in RPMI-C instead of saline and prefill the wells for this purpose first with 50µl of RPMI-C and then add 50µl of the spiking solution. The prefill for the unspiked testing of the dilution

is 100 µl of RPMI-C per well. Then add 50µl of the product dilution (or undiluted product, resp.) to 4 wells containing the spiking solution and 4 wells containing only RPMI-C.

Proceed according to the testing instructions (see above).
Determine the lowest dilution (highest concentration) of the product that yields an endotoxin spike recovery of 50 to 200%. For that purpose calculate the mean endotoxin values of the unspiked and the spiked product dilutions applying the 4-parameter logistic model on the endotoxin calibration curve. Consider that in this nomenclature "undiluted" means also a dilution. Then subtract the endotoxin value of the unspiked dilution from the endotoxin value of the corresponding spiked dilution. Calculate the spike recovery for each dilution in percent taking the theoretical value (spike concentration e.g. 0.25 EU/ml) as a 100%.
Example: Tested dilutions and spike recovery :"undiluted" = 25%; 1 in 2 = 49%; 1 in 4 = 90% and 1 in 8 = 110%. Then the dilution meeting the requirements is 1 in 4.

During the Validation Phase of the european study the test samples will be generally tested at the maximum valid dilution (MVD).

8.2.2 Interference with the ELISA system

Simulate a pyrogen test without cells and without incubation using the chosen test dilution of the product (from testing "interference with the cell system", in the example 1 in 4). Prepare enough replicates to spike with the IL-6 standard curve concentrations (4000 - 62.5 pg/ml and the blank) in duplicate (12 wells). Test the IL-6-spiked solutions directly for interference in the ELISA system together with a unspiked IL-6 standard dilution series. Apply exactly the same methodology as for a real pyrogen test. If interference in the ELISA system appears, the product has to be tested in further dilutions but not exceeding the maximum valid dilution (see above).
The lowest dilution (highest concentration) of the product not interfering with PBMC and the ELISA is considered as the future dilution for routine testing.

8.3 Prediction model

Perform the test with a certain product according to this SOP. Apply a valid test dilution showing no interference with the test systems as written above. Calculate the parameters of the endotoxin standard curve applying the 4-parameter logistic model. If this model does not fit your endotoxin concentrations, these have to be optimized and adapted to your local conditions.
Reject outliers only after checking according to Dixon's Test. Test on the level of p = 0.9 or 0.95.
Calculate the mean OD value of all replicates (at least triplicates). Calibrate the mean OD values of the product on the endotoxin standard curve and document the estimated endotoxin concentration. Multiply the estimated endotoxin concentration by the dilution factor (may be 1 in special cases). This value represents the pyrogenicity of the sample in terms of **endotoxin equivalents for the donor under test**. It should

not exceed the calculated endotoxin limit concentration. In this case the product shows a negative pyrogen reaction for this particular donor. Otherwise the sample is defined as showing a positive pyrogen reaction for that particular donor. Determine the reactions of all donors of the test. Since maximally 1 donor may be excluded from the test, there are either 4 or 3 reactions (from 4 or 3 donors) which are marked as "+" for positive and "-" for negative.

Decision level 1:

- If all donors (4 or 3) show a negative reaction, the product passes
- If 2 or more donors show a positive reaction, the product fails
- If 1 donor shows a positive reaction, an additional test with four different donors has to be performed (go to decision level 2)

Decision level 2

- If out of 6 to 8 donors a maximum of 1 donor show a positive reaction, the product passes.
- In any other case, the product fails

Remark: In order to optimize the prediction model of this test a methodology using a well characterized reference substance (identical in composition to the test substance) is in preparation. With this advanced prediction model interferences in the test due to donor variability will be reduced markedly.

9 HEALTH SAFETY AND ENVIRONMENT

Human material

Human material should be treated as biologically hazardous and all work using human material is to be carried out according to the procedures specified in the NOVARTIS Safety Guidelines.

Cultures of human material should be treated as biologically hazardous waste and disposed of according to the procedures specified in the NOVARTIS Safety Guidelines.

Bacterial endotoxin is, as its name indicates, a toxic agent and should be handled with care.

Precautions: Cover open cuts before use. Do not get in eyes, on skin, on clothing. Avoid inhaling. Keep container closed.

First Aid: In case of contact, immediately flush eyes or skin with plenty of water for at least 15 minutes. If inhaled, remove to fresh air. If not breathing, give artificial respiration, preferably mouth-to-mouth. If breathing is difficult, give oxygen.

Effects of skin absorption can include fever, headache and hypotension.

Effects of inhalation can include fever, headache and hypotension.

Effects of ingestion - adverse effects are unlikely since ingested endotoxin is rapidly detoxified.

10 ANNEX

11 REFERENCES

Bleeker, W.K., de Groot, E.M., den Boer, P.J., Biessels, P.T., Aarden, L.A. and Bakker, J.C. (1994) Measurement of interleukin-6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol. 22, 835-40.

Dinarello, C.A., Gatti, S. and Bartfai, T. (1999) Fever: links with an ancient receptor. Curr Biol. 9, R147-50.

Dinarello, C.A., O'Connor, J.V., LoPreste, G. and Swift, R.L. (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J. Clin. Microbiol. 20, 323-329.

Duff, G.W. and Atkins, E. (1982) The detection of endotoxin by *in vitro* production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52, 323-332.

Fennrich, S., Fischer, M., Hartung, T., Lexa, P., Montag-Lessing, T., Sonntag, H.G., Weigandt, M. and Wendel. (1999) A Detection of endotoxins and other pyrogens using human whole blood. Dev Biol Stand. 101, 131-9.

Finney DJ: Statistical Method in Biological Assay, Third Edition. London, Charles Griffin and Company Ltd, 1978.

Gaines Das RE and Tydeman MS. (1982) Iterative weighted regression analysis of logit responses: A computer program for analysis of bioassays and immunoassays. Computer Programs in Biomedicine 15; 13-22.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15, 425-33.

Levin, J. and Bang, F.B. (1964) A description of cellular coagulation in the Limulus. Bull. John Hopkins Hosp. 155, 337-345.

Mascoli, C.C. and Weary, M.E. (1979a) Limulus amebocyte lysate (LAL) test for detecting pyrogens in parenteral injectable products and medical devices: advantages to manufacturers and regulatory officials. J Parenter Drug Assoc. 33, 81-95.

Mascoli, C.C. and Weary, M.E. (1979b) Applications and advantages of the Limulus amebocyte lysate (LAL) pyrogen test for parenteral injectable products. Prog Clin Biol Res. 29, 387-402.

Meager, A., Parti, S., Leung, H., Peil, E. and Mahon, B. (1987) Preparation and characterization of monoclonal antibodies directed against antigenic determinants of recombinant human tumour necrosis factor (rTNF). Hybridoma 6, 305-311.

Poole, S., Selkirk, S., Rafferty, B., Meager, A., Thorpe, R. and Gearing, A. (1989) Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the

European Workshop on detection and quantification of pyrogen. Pharmeuropa, November 1989, 17-18.

Poole, S., Thorpe, R., Meager, A., Hubbard, A.R. and Gearing, A.J.H. (1988) Detection of pyrogen by cytokine release. Lancet 8577, 130.

Ray, A., Redhead, K., Selkirk, S. and Poole, S. (1990) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211-218.

Schins, R.P.F., van Hartingsveldt, B. and Borm P.J. (1996) *Ex vivo* cytokine release from whole blood. Exp. Toxic Pathol. 48, 494-496.

Taktak, Y.S., Selkirk, S., Bristow, A.F., Carpenter, A., Ball, C., Rafferty, B. and Poole, S. (1991) Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43, 578-582.

[This Page Intentionally Left Blank]

Appendix A5

The Monocytoid Cell Line Mono Mac 6 (MM6)/IL-6 *In Vitro* Pyrogen Test

ECVAM Background Review Document (March 2006) A-461
ECVAM Standard Operating Procedure (April 2002) A-529
ECVAM Standard Operating Procedure for the Validation Phase (October 2002) A-558

[This Page Intentionally Left Blank]

BRD: MM6/IL-6 March 2006

AN ALTERNATIVE IN VITRO PYROGEN TEST USING THE HUMAN MONOCYTOID CELL LINE MONO MAC-6 (MM6/IL-6)

BRD: MM6/IL-6 March, 2006

Contents

1 RATIONALE FOR THE PROPOSED TEST METHOD 4

1.1 INTRODUCTION 4
1.2 REGULATORY RATIONALE AND APPLICABILITY 5
1.3 SCIENTIFIC BASIS FOR THE PROPOSED TEST METHOD 6

2 TEST METHOD PROTOCOL COMPONENTS 8

2.1 OVERVIEW OF TEST METHOD 8
2.2 RATIONAL FOR SELECTED TEST COMPONENTS 9
2.3 BASIS FOR SELECTION OF THIS TEST METHOD 12
2.4 PROPRIETARY COMPONENTS 12
2.5 REPLICATES 13
2.6 MODIFICATIONS APPLIED AFTER VALIDATION 13
2.7 DIFFERENCES WITH SIMILAR TEST METHODS 14

3 SUBSTANCES USED FOR VALIDATION 15

3.1 SELECTION OF SUBSTANCES USED 15
3.2 NUMBER OF SUBSTANCES 15
3.3 DESCRIPTION OF SUBSTANCES USED 16
3.4 SAMPLE CODING PROCEDURE 17
3.5 RECOMMENDED REFERENCE CHEMICALS 17

4 *IN VIVO* REFERENCE DATA ON ACCURACY 18

4.1 TEST PROTOCOL *IN VIVO* REFERENCE TEST METHOD 18
4.2 ACCURACY 18
4.3 ORIGINAL RECORDS 19
4.4 QUALITY OF DATA 19
4.5 TOXICOLOGY 20
4.6 BACKGROUND ON ASSAY PERFORMANCE 20

5 TEST METHOD DATA AND RESULTS 21

5.1 TEST METHOD PROTOCOL 21
5.2 ACCURACY AND RELIABILITY 23
5.3 STATISTICS 31
5.4 TABULATED RESULTS 34
5.5 CODING OF DATA 36
5.6 CIRCUMSTANCES 36
5.7 OTHER DATA AVAILABLE 36

6 TEST METHOD ACCURACY 37

6.1 ACCURACY 37
6.2 CONCORDANCY TO *IN VIVO* REFERENCE METHOD 39
6.3 COMPARISON WITH REFERENCE METHODS 39
6.4 STRENGTH AND LIMITATIONS 39
6.5 DATA INTERPRETATION 40
6.6 COMPARISON TO OTHER METHODS 40

7 TEST METHOD RELIABILITY (REPEATABILITY/REPRODUCIBILITY) 41

7.1 SELECTION OF SUBSTANCES 41
7.2 RESULTS 41
7.3 HISTORICAL DATA 45
7.4 COMPARISON TO OTHER METHODS 45

8 TEST METHOD DATA QUALITY 46

8.1 CONFORMITY 46
8.2 AUDITS 46
8.3 DEVIATIONS 46
8.4 RAW DATA 46

9 OTHER SCIENTIFIC REPORTS AND REVIEWS 47

9.1 SUMMARY 47
9.2 DISCUSSION 48
9.3 RESULTS OF SIMILAR VALIDATED METHOD 48

10 ANIMAL WELFARE CONSIDERATIONS (REFINEMENT, REDUCTION, AND REPLACEMENT) 51

10.1 DIMINISH ANIMAL USE 51
10.2 CONTINUATION OF ANIMAL USE 51

11 PRACTICAL CONSIDERATIONS 53

11.1 TRANSFERABILITY 53
11.2 TRAINING 53
11.3 COST CONSIDERATIONS 54
11.4 TIME CONSIDERATIONS 54

12 REFERENCES 55

13 SUPPORTING MATERIALS (APPENDICES) 57

13.1 STANDARD OPERATING PROCEDURE (SOP) OF THE PROPOSED METHOD 57
13.2 STANDARD OPERATING PROCEDURE (SOP) OF THE REFERENCE METHOD 57
13.3 PUBLICATIONS 57
13.4 ORIGINAL DATA 60
13.5 PERFORMANCE STANDARDS 60

Appendix A – Method protocol(s) and trial plan(s)

Appendix B – Hardcopies of relevant publications

Appendix C – List of abbreviations and definitions

BRD: MM6/IL-6 March, 2006

1 Rationale for the Proposed Test Method

1.1 Introduction

1.1.1. *Describe the historical background for the proposed test method, from original concept to present. This should include the rationale for its development, an overview of prior development and validation activities, and, if applicable, the extent to which the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards.*

Pyrogens are a chemically heterogeneous group of hyperthermia- or fever-inducing compounds. They derive from bacteria, viruses, fungi or from the host himself. A subject reacts to microbial products (exogenous pyrogens) by producing endogenous pyrogens such as prostaglandins and the pro-inflammatory cytokines interleukin-1 (IL-1), interleukin-6 (IL-6) and tumour necrosis factor-α (TNF-α) (Dinarello, 1999). Depending on the type and amount of pyrogen challenge and the sensitivity of an individual, even life-threatening shock-like conditions can be provoked. To assure quality and safety of any pharmaceutical product for parenteral application in humans, pyrogen testing is therefore imperative.

Depending on the medicinal product, one of two animal-based pyrogen tests is currently prescribed by the respective Pharmacopoeias, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET). For the rabbit pyrogen test, sterile test substances are injected intravenously to rabbits and any rise in body temperature is assessed. This *in vivo* test detects various pyrogens but not alone the fact that large numbers of animals are required to identify a few batches of pyrogen-containing samples argues against its use. In the past two decades, the declared intention to refine, reduce and replace animal testing, has lowered rabbit pyrogen testing by 80% by allowing to use the BET as an *in vitro* alternative pyrogen test for certain medicinal products (Cooper et al, 1971).

Bacterial endotoxin, comprising largely lipopolysaccharide (LPS) from the cell wall of Gram-negative bacteria that stimulates monocytes/macrophages via interaction with CD14 and toll-like receptor 4 (TLR4) (Beutler and Rietschel, 2003), is the pyrogen of major concern to the pharmaceutical industry due to its ubiquitous sources, its stability and its high pyrogenicity. With the BET, endotoxin is detected by its capacity to coagulate the amoebocyte lysate from the haemolymph of the American horseshoe crab, *Limulus polyphemus*, a principle recognised some 40 years ago (Levin and Bang, 1964). In the US, *Limulus* crabs are generally released into nature after drawing about 20% of their blood and therefore most of these animals survive. However, the procedure still causes mortality of about 30,000 horseshoe crabs per year, which adds to even more efficient threats of the horseshoe crab population like its use as bait for fisheries, habitat loss and pollution (http://www.horseshoecrab.org/).

As with the rabbit test the general problem of translation of the test results to the human fever reaction persists. Moreover, although being highly sensitive, the failure of the BET to detect non-endotoxin pyrogens as well as its susceptibility to interference by e.g. high protein or lipid levels of test substances or by glucans impedes full replacement of the rabbit pyrogen test. Hence, hundreds-of-thousands rabbits per year are still used for pyrogen testing.

A test system that combines the high sensitivity and *in vitro* performance of the BET test with the wide range of pyrogens detectable by the rabbit pyrogen test is therefore required in order to close the current testing gap for pyrogen and to avoid animal-based tests. With this intention and due to improved understanding of the human fever reaction (Dinarello, 1999), test systems based on *in vitro* activation of human monocytes were developed. First efforts date back about 20 years, when peripheral blood mononuclear cells (PBMC) were used to detect endotoxin by monitoring the release of pyrogenic cytokines (Duff and Atkins, 1982; Dinarello et al, 1984). Meanwhile, a number of different test systems, using either whole blood, peripheral blood mononuclear cells (PBMCs) or the monocytoid cell lines MONO MAC 6 (MM6) or THP-1 as a source for human monocytes and various read-outs were established (Poole et al., 1988; Ziegler et al, 1988; Tsuchiya et al, 1980; Hartung and Wendel, 1996; Hartung et al, 2001; Poole et al, 2003). These test systems were validated with the aim of developing a tool for formal inclusion into Pharmacopoeias, an important basis for implementing novel alternative pyrogen tests for product-specific validation.

1.1.2 *Summarize and provide the results of any peer review conducted to date and summarize any ongoing or planned reviews.*
All of the five methods are currently under peer review of the ECVAM Scientific Advisory Committee.

1.1.3 *Clearly indicate any confidential information associated with the test method; however, the inclusion of confidential information is discouraged.*
This document does not contain any confidential information.

1.2 Regulatory rationale and applicability

1.2.1 *Describe the current regulatory testing requirement(s) for which the proposed test method is applicable.*
To assure quality and safety of pharmaceutical products for parenteral application in humans, pyrogen testing is imperative. Depending on the drug, one of two pyrogen tests is currently prescribed by the European Pharmacopoeia, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET), and other national and international guidelines.

1.2.2 *Describe the intended regulatory use(s) (e.g., screen, substitute, replacement, or adjunct) of the proposed test method and how it will be used to substitute, replace, or complement any existing regulatory testing requirement(s).*
Dependent on the product and the presence of relevant clinical data on unexpected pyrogenicity of clinical lots, the proposed test method may be an alternative method for pyrogen testing, thus substituting the rabbit pyrogen test or the BET. In certain cases, the proposed test method may function as a supplementary test method to assess compliance to the licensing dossier.
In case the proposed test method is an alternative for pyrogenicity testing, a thorough cross-validation between the proposed test method and the original method for the specific medicinal product is warranted. In case the proposed test method is an adjunctive test to screen for (unexpected) pyrogenic lots, alert and alarm limits may be established based on consistency of production lots or (preferably) based on actual clinical data.

1.2.3 *Where applicable, discuss the similarities and differences in the endpoint measured in the proposed test method and the currently used in vivo reference test method and, if appropriate, between the proposed test method and a comparable validated test method with established performance standards.*
The current *in vivo* method (rabbit test), as described in the pharmacopoeia, and the proposed *in vitro* test method each determine very different end-points, though the biochemical origins of the response are similar.
The *in vivo* method more resembles a black box, and determines the total rise in body temperature (fever induction) of the animals subjected to the medicinal product, as a result of pyrogens (if any) present in the product.
The proposed MM6/IL-6 test method is an *in vitro* model for the fever response mechanism. It determines the release of cytokines by monocytoid cells into the culture medium upon the interaction of pyrogens and specific receptors on the monocytoid cells. It is these cytokines that trigger the fever response *in vivo*.
Main difference between the *in vivo* and *in vitro* methods are that the latter is quantitative and uses cells of human origin, thus better reflecting the physiological situation.

1.2.4 *Describe how the proposed test method fits into the overall strategy of hazard or safety assessment. If a component of a tiered assessment process, indicate the weight that should be applied relative to other measures.*
The proposed MM6/IL-6 test method may be applied for those medicinal products for which the rabbit test is the only or most reliable method for pyrogenicity testing, since a) the medicinal product is not compatible with the BET or b) the medicinal product contains pyrogens other than Gram-negative endotoxin.
Limit concentrations for pyrogens are established based on consistency lots or actual clinical data or, in the case of endotoxin the endotoxin limit concentration (ELC) as defined for many medicinal products.

1.3 Scientific basis for the proposed test method

1.3.1 *Describe the purpose and mechanistic basis of the proposed test method.*
The proposed *in vitro* method is intended to determine the presence of pyrogens in medicinal products for parenteral use. The proposed test method is an *in vitro* model of the human fever response. It determines the release of cytokines (IL-6) upon the interaction of pyrogens and specific Toll-like receptors on the monocytoid cells (MM6) (Beutler and Rietschel, 2003). It is these cytokines that trigger the fever response *in vivo*.

1.3.2 *Describe what is known and not known about the similarities and differences of modes and mechanisms of action in the proposed test method as compared to the species of interest (e.g., humans for human health-related toxicity testing).*
An important feature of the proposed test method is that it is based upon the use of monocytoid cells of human origin. It therefore by definition resembles more closely the actual response of humans. The two other test methods make use of either crustaceans (BET) or rabbits, both species more or less distinct from the human species. The response of humans, horseshoe crabs and rabbits toward Gram-negative endotoxin has been studied extensively and the methods appear equivalent for this particular pyrogen

(Cooper et al 1971; Greisman and Hornick, 1969). However, there are documented cases of medicinal products and specified pyrogenic substances that yield false-positive or false-negative results in either test method. Since the proposed MM6/IL-6 test method is based on human monocytoid cells, it may therefore predict more accurately the pyrogenicity of such substances in humans.

1.3.3 *Describe the intended range of substances amenable to the proposed test method and/or the limits of the proposed test method according to chemical class or physicochemical factors.*

The proposed MM6/IL-6 test method is intended for the assessment of pyrogens in all parenteral medicinal products for human use, chemical or biological and including raw materials, bulk ingredients and excipients. Use of the proposed test method in testing environmental samples or medicinal products is suggested and may be feasible, but substantiating data are as yet limited or absent.

2 Test Method Protocol Components

2.1 Overview of test method.

Provide an overview of how the proposed test method is conducted. If appropriate, this would include the extent to which the protocol for the proposed test method adheres to established performance standards.

A highly detailed protocol of the proposed test method (*Detailed protocol MM6/IL-6: "In vitro pyrogen test using MONOMAC 6 CELLS"; electronic file name: SOP MM6-IL-6*) is attached in Appendix A of this background review document (BRD). Appendix A also includes the amended protocol used in the formal validation study to determine the sensitivity and specificity of the test (section 3, table 3.3.1). However, it does only replace the previous version for testing of parenteral drugs described in table 3.3.1, and was included into Appendix A for completeness of information only (*marked with internal identifier SopMM6v08; electronic file name SOP MM6-IL-6 validation)*.

To facilitate routine testing a continuous culture of the MM6 cells is maintained. Detailed information for starting up a cell culture, propagation of the cells and preparation of a cell bank is given in the method protocol in Appendix A of this BRD. Cell-culturing as well as a significant part of the test is carried out in a Class 2 laminar flow sterile cabinet using aseptic technique, and reagents and consumables that are sterile and pyrogen-free. Throughout the test there is no need for sophisticated or dedicated laboratory equipment. The procedures require only equipment that is readily available in a QC cell-culturing laboratory.

The test protocol itself can be divided into several sequential steps. First, samples to be tested are diluted to the appropriate concentration and four replicated are added to a 96-well cell culture plate. For reference a standardized dilution series of LPS is included. Second, a standardized number of MM6 cells from the cell culture is added to each well and incubated for 16-24 hours in a CO_2-incubator (37°C, 5% CO_2, high humidity). From each well an aliquot of the supernatant above the MM6-cells is transferred to the corresponding well in a new 96-wells cell culture plate. Subsequently, the concentration of IL-6 released in the cell-conditioned medium is quantified using a validated IL-6-specific ELISA, in which the International standard for IL-6 (WHO code: 89/548) or an IL-6 standard calibrated against the International Standard (IS) may be used as the assay calibrant.

The WHO-LPS standard (vial code 94/580, E.coli O113:H10:K-), was used throughout the validation. This standard is identical to USP Reference Standard Endotoxin (EC6). There are several possibilities to estimate the pyrogenic contamination of the preparations under test: 1) By the construction of a dose-response curve for endotoxin standard versus OD-value of the IL-6 ELISA. The contamination of the preparations is measured in endotoxin–equivalent units. 2) The inclusion of an endotoxin threshold control (e.g. one dilution of the standard curve) which allows for the classification in positive and negative

samples (i.e. pyrogenic and non-pyrogenic samples). 3) Inclusion of an appropriate positive product control.

2.2 Rational for selected test components

Provide a detailed description and rationale, if appropriate, for the following aspects of the proposed test method:

2.2.1 Materials, equipment, and supplies needed.

The materials, equipment and supplies used for the MM6/IL-6 test are laboratory items that will be already available in routine cell culturing laboratory. The materials that will be in close contact with samples and MM6-cells need to be pyrogen free. The items are specified in the method protocol enclosed in Appendix A. There are a few materials which will not be readily available and will need preparation: A working cell bank should be prepared with MM6-cells obtained from the NIBSC or from the DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, Germany.
The IL-6 ELISA can be either commercially obtained or hand made. Sensitivity and specificity should be established before use. Also the fetal bovine serum used during cell culturing should be tested for interference and pyrogenicity. In general, the likely sources of pyrogens are the culture medium and fetal bovine serum. Therefore pyrogen free culture medium should be purchased. The pyrogen content of the fetal bovine serum should be <0.1ng/ml (after heating for 30 min. at 70°C), as confirmed by the currently proposed test method or the BET.

2.2.2 Dose-selection procedures, including the need for any dose range-finding studies or acute toxicity data prior to conducting a study, if applicable.

For every kind of test compound the interference in the test cell culture and the test readout system is determined. For this purpose, a preliminary “dose finding” test is conducted to establish a suitable (interference free) dilution for every new test compound. For the validation study (as described in section 4of this BRD), the tested products were diluted according to their known ELC, which was usually far beyond interfering concentrations. The ELCs of the tested products or drugs were calculated according to the European Pharmacopoeia.

2.2.3 Endpoint(s) measured.

The proposed test method is an *in vitro* model of the fever response mechanism. It determines the release of interleukin 6 (IL-6) by the monocytoid cell line MM6. IL-6 is released into the culture medium upon the interaction of pyrogens and specific receptors on the monocytoid cells. The measured endpoint IL-6 is one of the cytokines that trigger the fever response *in vivo*.

2.2.4 Duration of exposure.

The MM6-cells are exposed to the potentially pyrogenic components in samples during 16-24 hours: After addition of the samples to be assayed to the MM6-cells, the cultures are subsequently incubated at 37°C for 16-24 hours in an atmosphere of 5% CO2 in humidified air. The exposure is discontinued by transferring the supernatant above the MM6-cells to a new 96-wells cell culture plate.

2.2.5 Known limits of use.
The method as described in the protocol is not a finalized test system for the testing of all medicinal products. The method may be applied only to preparations that have been validated with the method, i.e. shown not to interfere in the cell culture and the readout system a specified dilution. A section describing the interference testing is included in the protocol (see Annex I). However, at this moment there are no medicinal products known that can not be tested with the method.

2.2.6 Nature of the response assessed.
The proposed MM6/IL-6 test method is an *in vitro* model of the fever response mechanism. Upon the interaction of exogenous pyrogens and specific receptors on the monocytoid cells endogenous pyrogens (e.g. interleukins, TNF-□ and prostaglandins) are produced. In the body the fever response is triggered by these endogenous pyrogens. Immunoreactive IL-6, the measured endpoint for the current method, is one of these endogenous pyrogens. The human monocytoid cell line MM6 serves as a model for the different types of monocytoid cells in the blood.

2.2.7 Appropriate vehicle, positive, and negative controls and the basis for their selection.
Throughout the development and validation phase the test compounds are diluted in 0.9% clinical saline. This 0.9% clinical saline is considered an appropriate vehicle as no interference with active substances of a drug is to be expected.
In addition the test includes several controls.
A negative control: 0.9% clinical saline (sodium chloride)
A positive control: WHO-LPS 94/580, 0.5 EU/ml in clinical saline.
A negative product control: clean, released batch for each drug.
A positive product control: test item spike with WHO-LPS (code 94/580) at 0.5 EU/ml
The positive and negative controls are the same in every assay and are needed to establish the sensitivity of the test system. In addition, a product-based set of controls is used to reveal product-related interference.

2.2.8 Acceptable range of vehicle, positive and negative control responses and the basis for the acceptable ranges.
An MM6/IL-6 test is considered acceptable for further analysis if the positive control (0.5 EU/ml) is significant elevated over the negative control (0.9% clinical saline). Moreover the response to different concentrations of the positive control should show a dose response relationship. To be able to quantify the responses to the positive control this should be well within the maximum response that can be measured with the test system.

As regards the substances to be tested, for products with an established ELC, specified in EU/ml, the product is diluted to its maximum valid dilution (MVD). The negative product control should be negative at the MVD. The response to the positive product control should be between 50% and 200% of the response to the positive control, indicating a possible pyrogenicity can be detected using these conditions.

2.2.9 Nature of the data to be collected and the methods used for data collection.

The raw data collected are the read-outs (Absorbance) of the IL-6 ELISA, measured by an automated laboratory ELISA-plate reader. The wavelength is dependent on the chromogenic substrate applied.

2.2.10 Type of media in which data are stored.
Data are stored in electronic files (windows98 compatible software) and as hard copy.

2.2.11 Measures of variability.
As part of the development of the MM6/IL-6 test the intralaboratory repeatability was assessed by independent and identical replicated measurement of the different concentrations of WHO-LPS. Furthermore, the limit of detection and its dependence from known but uncontrollable variables such us operator and passage of the cell line were investigated. These variables and the inherent variation of biological systems make up to the total variation of the method.

2.2.12 Statistical or non-statistical methods used to analyze the resulting data, including methods to analyze for a dose-response relationship. Justify and describe the method(s) employed.
All experiments are run with four replicates of the test compound on one plate. A standard curve in quadruplicate, using the International Standard for Endotoxin (calibrated in EU) is included, ranging from 0.25 EU/ml up to 4 EU/ml (or 0.125 EU/ml up to 2 EU/ml). Outliers are rejected only after checking according to the Dixon's test, which is USP approved, and applied to identify and eliminate aberrant data. Next, the negative and the respective positive control are compared to ensure a suitable limit of detection. For this a one sided t-test is applied to the log transformed data to ensure that the response to the positive control is significantly larger than that of the respective negative control.
The endotoxin value of each replicate calculated from the endotoxin calibration curve of the IS for endotoxin, applying the 4-parameter logistic model and expressed as endotoxin-equivalents/ml (EU/ml). Subsequently, the mean endotoxin value of all replicates (usually quadruplicates) of a test compound is calculated and multiplied by the dilution factor (if applicable).

2.2.13 Decision criteria and the basis for the prediction model used to classify a test chemical (e.g., positive, negative, or equivocal), as appropriate.
A PM was developed in order to classify substances as "pyrogenic for humans" or "non-pyrogenic for humans". To be able to define a dichotome result in the alternative pyrogen test, a threshold pyrogen value of 0.5 EU/ml was chosen. This threshold value was based on historical data with rabbits (described in section 4.1). The suitability of the PM was assessed by testing substances which were artificially contaminated with endotoxin (substances are described in section 3.2 and 3.3). The statistical approach, including quality criteria, is detailed in section 5.3.

2.2.14 Information and data that will be included in the study report and availability of standard forms for data collection and submission.

Raw data were collected using a standard form. These were submitted to the quality department of ECVAM.

2.3 Basis for selection of this test method

Explain the basis for selection of the test method system. If an animal model is being used, this should include the rationale for selecting the species, strain or stock, sex, acceptable age range, diet, and other applicable parameters.

In view of the shortcomings of the rabbit pyrogen test and the BET, *in vitro* pyrogen tests that utilise the exquisite sensitivity to exogenous pyrogen of monocytoids have been proposed. In such tests, products are incubated with human peripheral blood monocytes (or mononuclear cells, PBMC cells or leukocytes) and the conditioned media assayed for pyrogenic cytokines (Duff & Atkins, 1982; Dinarello et al., 1984; Poole et al., 1988; Poole et al, 1989; Hansen and Christensen, 1990; Taktak et al., 1991; Bleeker et al., 1994).

The isolation of monocytes/leukocytes from whole blood is labor–intensive and time–consuming, technically sophisticated, requires expensive reagents and does not guarantee the isolation of cells in a non–activated state. This prompted the evaluation of various cell lines which retain monocytoid characteristics, including the capacity to synthesize and secrete pyrogenic cytokines (Taktak et al., 1991). Taktak *et al*. measured LPS–induced IL–1β and IL–6 release from MONO MAC-6 (MM6) cells and THP–1 cells and concluded that IL–6 release by MM6 cells was the most appropriate readout for an *in vitro* pyrogen test ('monocyte test') because immunoreactive IL–6, unlike immunoreactive IL–1 and TNFα, is secreted entirely into the cell–conditioned medium in large quantities, permitting its complete estimation.

2.4 Proprietary components

If the test method employs proprietary components, describe what procedures are used to ensure their integrity (in terms of reliability and accuracy) from "lot-to-lot" and over time. Also describe procedures that the user may employ to verify the integrity of the proprietary components.

The MM6 cell line was established by Prof. H.W.L. Ziegler-Heitbrock, Institute for Immunology, Universtiy of Munich, Munich, Germany. The cell line can be obtained for research purposes only from Prof. Ziegler-Heitbrock or from the German Collection of Microorganisms and Cell Cultures (DSMZ) Braunschweig, Germany. The conditions for licensing of the cell line are to be negotiated individually with Prof.Ziegler-Heitbrock.

A Master Cell Bank (MCB) and a Working Cell Bank (WCB) were established at the NIBSC. MM6 cells can be obtained from this laboratory. In addition, they can be obtained from DSMZ, Braunschweig, Germany. However, the performance of batches from these sources has not been compared with the NIBSC MCB and WCB used during the present study. The reliability and accuracy of the MM6 cells used for each test is ensured by including appropriate positive and negative controls. As a positive control a specified amount of Units of the International Endotoxin Standard is used and test criteria should be met. Minimum requirements are set for variability of replicates.

The response to the negative control (usually cell-culture medium) should be well below detection level.

2.5 Replicates

Describe the basis for the number of replicate and repeat experiments; provide the rationale if experiments are not replicated or repeated.

All experiments are run with four replicates of the test compound on one plate. A standard curve in quadruplicate, using the International Standard for Endotoxin (calibrated in EU) is included, ranging from 0.25 EU/ml up to 4 EU/ml (or 0.125 EU/ml up to 2 EU/ml). A test is valid if the 0.5 EU/ml is significant elevated over the background (defined by mean of the blank values +2SD (n-1) of the blank values). The endotoxin value of each replicate is calculated from the endotoxin calibration curve of the IS for endotoxin, applying the 4-parameter logistic model and expressed as endotoxin-equivalents/ml (EU/ml). Outliers are rejected only after checking according to the Dixon's test ($p \geq 0.05$). Four replicates is considered the minimal amount for the Dixon's test.

During a prevalidation (PV) phase, the intralaboratory reproducibility as well as the interlaboratory reproducibility of the MM6/IL6 test were established by applying repeated experiments (detailed in section 7). As the test method reliability (repeatability/reproducibility) was shown to be satisfactory, it was feasible to perform the test with pharmaceutical substances (detailed in table 1) once by the three participating laboratories.

2.6 Modifications applied after validation

Discuss the basis for any modifications to the proposed test method protocol that were made based on results from validation studies.

During the prevalidation study it was shown that pre-incubation of the cells, comprising of adjusting the cells to a specified number per ml and replacing the culture-medium, highly improved the response and the repeatability of the test. Although this extends the timeframe of the test by one day, this procedure was worth adopting for the MM6/IL6 test and was subsequently been used throughout the validation phase.

The initial protocol allowed the use of various kinds of IL-6 ELISAs, however, due to their sub-optimal repeatability their use was restricted to the two IL-6 ELISAs now indicated in the protocol (the in-house Novartis IL-6 ELISA and the CLB Human IL-6 ELISA kit). Both IL-6 ELISAs use the same monoclonal antibody for IL-6 detection. It should be noted that these ELISAs may be substituted with other validated IL-6 specific ELISAs, in which the International Standard (IS) for IL-6 (or an IL-6 standard calibrated against the IS) is used as the assay calibrant.

The most significant modification to the test method protocol, concerns the **prediction model.** For the purpose of the validation study, a sample is considered positive when the mean OD at 450 nm to the tested product exceeds the mean OD at 450 nm of that of the positive product control.

2.7 Differences with similar test methods

If applicable, discuss any differences between the protocol for the proposed test method and that for a comparable validated test method with established performance standards.
Not applicable.

3 Substances Used for Validation

3.1 Selection of substances used

Describe the rationale for the chemicals or products selected for use in the validation process. Include information on the suitability of the substances selected for testing, indicating any chemicals that were found to be unsuitable.

Selected test items were medicinal products available on the market. Released clinical batches were considered clean, i.e. containing no detectable pyrogens. To test the specificity, sensitivity and the reproducibility of the proposed test method, the products were spiked with pyrogen. For the present studies endotoxin (LPS) was selected as the model pyrogen, since it is well defined, standardized and readily available.

For the sensitivity and specificity the test items were assessed at their MVD. The MVD is the quotient of the ELC and the detection limit. The European Pharmacopoeia prescribes for various types of parenterals the amount of endotoxin that is maximally allowed in a medicinal product, i.e. the ELC, taking into consideration the dose, the route of administration and the dosing regimen of the product.

The aim of the study was to discriminate between negative and positive samples. Based on the selected pyrogen threshold value (see 4.1) and the PM applied, positive samples were defined as containing 0.5 EU/ml or more. Hence, to determine the MVD, the value of 0.5 EU/ml was defined as the detection limit.

Test items were assessed as such (negative product control), spiked with endotoxin at 0.5 IU/ml (positive product control) and after spiking with endotoxin at 5 levels (blinded samples). In addition a negative control (saline) and positive control (0.5 EU/ml in saline) were included to establish assay validity.

For reproducibility, the test items were tested, at a predefined dilution above the MVD, independently in 3 different laboratories, 3 times each. Based on the selected pyrogen threshold value (see 4.1) and the PM applied, positive samples were defined as containing 0.5 EU/ml or more. The test items were tested after spiking with endotoxin at four levels. For no other reasons but practical ones, i.e. availability of test materials, different test items were selected for this part of the validation study.

It was determined earlier whether candidate test items interfered with the outcome of the proposed test method. Interference was considered when the response of endotoxin in the diluted test item was below 50% or above 200% of the response of endotoxin in saline (spike-recovery). It was shown that none of the test items interfered with the assay at the selected dilutions (data not shown).

3.2 Number of substances

Discuss the rationale for the number of substances that were tested.

A total of 13 test items were selected for the validation study (see 3.3): 10 test items for determining sensitivity and specificity (table 3.3.1), 3 different test items for determining reproducibility (table 3.3.2). Test items and their spikes were appropriately blinded by ECVAM before distribution to the participating testing facilities.

For sensitivity and specificity, each test item was tested after spiking at its individual MVD. Hence they each came with their own specific set of 5 endotoxin spike solutions: 0.0, 0.25, 0.5, 0.5 and 1.0 EU/ml. Simple logistics limited the amount of test items for this part of the validation study to 10. Since test items were assessed with 5 different endotoxin levels at 3 independent laboratories, this yielded a total of 150 data points, biometrically considered to be sufficient for further analysis.

For reproducibility each test item was spiked at 4 different levels (0.0, 0.0, 0.5 and 1.0 EU/ml) and tested at specified dilutions, 3 times in 3 laboratories.

3.3 Description of substances used

Table 3.3.1: Test items (parenteral drugs) used for determining sensitivity and specificity

Drug	code	Source	Agent	Indication	MVD (-fold)
Glucose 5% (w/v)	GL	Eifel	Glucose	nutrition	70
Ethanol 13% (w/w)	ET	B.Braun	Ethanol	diluent	35
MCP®	ME	Hexal	Metoclopramid	antiemetic	350
Orasthin®	OR	Aventis	Oxytocin	initiation of delivery	700
Binotal®	BI	Aventis	Ampicillin	antibiotic	140
Fenistil®	FE	Novartis	Dimetindenmaleat	antiallergic	175
Sostril®	SO	GlaxoSmithKline	Ranitidine	antiacidic	140
Beloc®	BE	Astra Zeneca	Metoprolol tartrate	heart dysfunction	140
Drug A*	LO	-	0.9% NaCl	-	35
Drug B*	MO	-	0.9% NaCl	-	70

*Drugs A and B were included as saline controls using notional ELCs.
Negative control: 0.9% clinical stock saline solution.
Positive control: WHO-LPS 94/580 (*E. coli* 0113:H10:K-), 0.5 EU/ml in clinical stock saline.

Table 3.3.2: Test items (parenteral drugs) used for determining reproducibility.

Drug	Source	Agent	Indication
Gelafundin®	Braun melsungen	Gelatin	Transfusion
Jonosteril ®	Fresenius	Electrolytes	Infusion
Haemate ®	Aventis	Factor VIII	Hemophilia

Negative control: 0.9% clinical stock saline solution.
Positive control: WHO-LPS 94/580 (*E. coli* 0113:H10:K-), 0.5 EU/ml in clinical stock saline.

3.4 Sample coding procedure

Describe the coding procedures used in the validation studies.
All test items are registered medicinal products and were obtained from a pharmaceutical supplier. Test items and endotoxin spiking samples were prepared, blinded where appropriate and coded under GLP by personnel from ECVAM, Italy. These were then taken over by the Paul-Ehrlich Institute, Germany, for allocation and shipment to each of the appropriate test facilities participating in the study.
For the sensitivity and specificity part of this study, test items and their respective spikes (5 per test item) were all blinded. For reproducibility testing, only the spikes (4) were blinded, the test items were not.

3.5 Recommended reference chemicals

For proposed test methods that are mechanistically and functionally similar to a validated test method with established performance standards, discuss the extent to which the recommended reference chemicals were tested in the proposed test method. In situations where a listed reference chemical was unavailable, the criteria used to select a replacement chemical should be described. To the extent possible, when compared to the original reference chemical, the replacement chemical should be from the same chemical/product class and produce similar effects in the in vivo reference test method. In addition, if applicable, the replacement chemical should have been tested in the mechanistically and functionally similar validated test method. If applicable, the rationale for adding additional chemicals and the adequacy of data from the in vivo reference test method or the species of interest should be provided.
The reference pyrogen material used was the international endotoxin standard WHO-LPS 94/580 (*E. coli* 0113:H10:K-). Where appropriate, the material was diluted in clinical saline solution (0.9% (w/v) sodium chloride). The saline was also used as negative control (blank).

4 *In vivo* Reference Data on Accuracy

4.1 Test protocol *in vivo* reference test method.

Provide a clear description of the protocol(s) used to generate data from the in vivo reference test method. If a specific guideline has been followed, it should be provided. Any deviations should be indicated, including the rationale for the deviation.
For ethical reasons, no rabbit pyrogen tests were performed for this study. However, Dr. U. Lüderitz-Püchel, Paul-Ehrlich Institute, Germany, kindly provided historical data, accumulated over several years, from 171 rabbits (Chinchilla Bastards). The respective Pharmacopoeiae do not prescribe a rabbit strain for the *in vivo* pyrogen test, but Chinchilla rabbits are reported as a relatively sensitive strain for pyrogen testing.

The rabbits were injected with endotoxin and their rise in body temperature over the next 180 minutes was recorded (figure 4.1.1). From these data it was established that 50% of the rabbits got fever when treated with endotoxin at 5 EU/kg (Hoffmann et al, 2005a). Fever in rabbits is defined as a rise in body temperature over 0.55°C. On the basis of these historical animal data and corrected for the maximal volume allowed in rabbits, i.e. 10 ml per animal, a pyrogen threshold value of 0.5 EU/ml was defined for the PM in the proposed test method.

4.2 Accuracy

Provide the in vivo reference test method data used to assess the accuracy of the proposed test method. Individual human and/or animal reference test data, if available, should be provided. Provide the source of the reference data, including the literature citation for published data, or the laboratory study director and year generated for unpublished data.
As mentioned, animal studies were not performed due to ethical reasons. However, a theoretical assumption on specificity and sensitivity can be made (Hoffmann et al, 2005a). Taking the prevalence of the different final concentrations of LPS in the 5 spikes into account (1.0 EU/ml: 20%; 0.5 EU/ml: 40%; 0.25 EU/ml: 20% and 0.0 EU/ml: 20%) and calculating probabilities of misclassification using the chosen study design and defined threshold of pyrogenicity, i.e. 0.5 EU/ml, the theoretical sensitivity of the rabbit pyrogen test is 57.9% and the theoretical specificity of the rabbit pyrogen test is 88.3%.

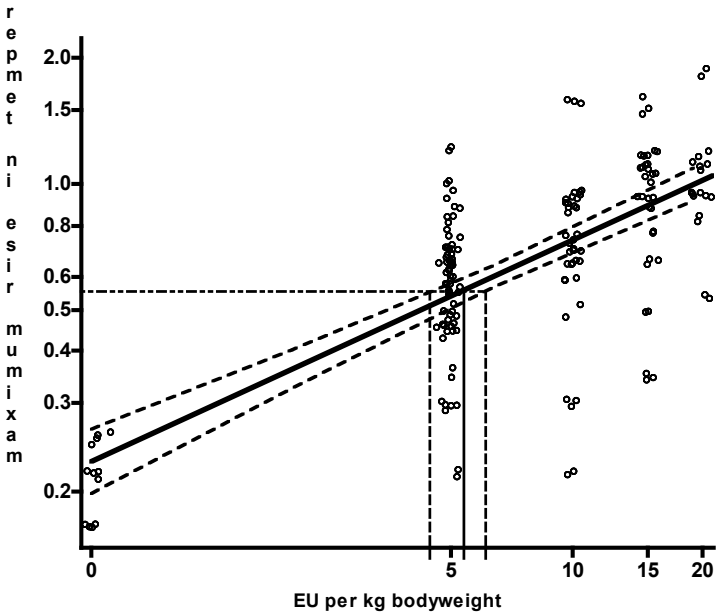


Figure 4.1.1 Dose-temperature of standard endotoxin applied to Chinchilla Bastards (n=171). Rabbits were treated with 1 ml saline containing 0, 5, 10, 15 and 20 EU of *E. coli* LPS (WHO-LPS 94/580 (E.coli O113:H10:K)) and their body temperature was measured over 180 min. Linear regression analysis was performed after logarithmic transformation of the data. Data are shown as dots to which a jitter-effect was applied in order to be able to distinguish congruent data. The full line depicts the linear regression whereas the dashed lines represent the 95%-confidence bounds. Furthermore, a horizontal line for a 0.55°C raise of temperature is added which is often defined as the rabbit threshold for fever. At the interception point of this line and the regression line 50% of the rabbits are to be expected to develop fever.

4.3 Original records

If not included in the submission, indicate if original records are available for the in vivo reference test method data.

The recognition of pyrogenic substances as bacterial by-products and the identification of a variety of pyrogenic agents enabled the development of a proper test to demonstrate non-pyrogenicity of the pharmaceutical product. As early as the 1920s, studies were done to select the most appropriate animal model. Results indicated that most mammals had a pyrogenic response, but only a few, including rabbits, dogs, cats, monkeys and horses showed a response similar to that in humans. For practical reasons, other species but rabbits and dogs were considered not practical. Co Tui & Schrift (1942) described that rabbits are less thermo-stable than dogs. Hence, rabbits are more suited for the purpose of testing for the absence of pyrogens, since a negative result is more significant.

4.4 Quality of data

Indicate the quality of the in vivo reference test method data, including the extent of GLP compliance and any use of coded chemicals.

The study protocol(s) and all standard operating procedures were GLP-concordant. These documents were quality assured by the ECVAM quality assurance officers.

4.5 Toxicology

Discuss the availability and use of relevant toxicity information from the species of interest (e.g., human studies and reported toxicity from accidental or occupational exposure for human health-related toxicity testing).

Over time, a number of studies were performed to correlate the rabbit test to pyrogenic reactions in humans. A conclusive study by Greisman and Hornick, published in 1969, who compared three purified endotoxin preparations (*Salmonella typhosa*, *E. Coli* and *Pseudomonas*) in New Zealand rabbits and in male volunteers, showed that the induction of a threshold pyrogenic response, on a weight basis, was similar to rabbit and man. At higher doses, rabbits respond less severe as compared to man.

4.6 Background on assay performance

Discuss what is known or not known about the accuracy and reliability of the in vivo reference test method.

As mentioned, animal studies were not performed due to ethical reasons. However, a theoretical assumption on specificity and sensitivity can be made (Hoffmann et al, 2005a). Taking the prevalence of the different final concentrations of LPS in the 5 spikes into account (1.0 EU/ml: 20%; 0.5 EU/ml: 40%; 0.25 EU/ml: 20% and 0.0 EU/ml: 20%) and calculating probabilities of misclassification using the chosen study design and defined threshold of pyrogenicity, i.e. 0.5 EU/ml, the theoretical sensitivity of the rabbit pyrogen test is 57.9% and the theoretical specificity of the rabbit pyrogen test is 88.3%.

5 Test Method Data and Results

5.1 Test method protocol

Describe the proposed test method protocol used to generate each submitted set of data. Any differences from the proposed test method protocol should be described, and a rationale or explanation for the difference provided. Any protocol modifications made during the development process and their impact should be clearly stated for each data set.

The protocol for the MM6/IL-6 test is provided in Appendix A of this BRD. It includes the precise step-by-step description of the test method and lists the necessary reagents and laboratory procedures for generating data. For two steps during validation, a part of the protocol was adapted to contain a detailed description of the dilution of the samples and the spiking with WHO-LPS. The relevant part of the protocol is detailed in this section as well. The protocols are attached in the Appendix A for the sake of completeness of the documentation only. The validity criteria and the detailed statistical analysis described in section 5.3 of this BRD were applied to analyse the data produced during validation.

To assess the reliability of the test method a series of experiments were conducted in one laboratory (the developing laboratory [DL]). As a start, only blanks (saline, 0 EU/ml) and spikes of WHO-LPS in saline were tested. These experiments are summarised in table 5.1.1.

Table 5.1.1: summary of experiments with WHO-LPS in saline

Experiment	Spikes (EU/ml) in saline	n (per spike)	Repetitions of experiment	N
1A	0; 0.25; 0.5	20	1	60
1B	0; 0.063; 0.125; 0.25; 0;5	12	1	60
2A	0; 0.25; 0.5	20	3	180
2B	0; 0.25; 0.5	20	3	180

The collected data were used to answer questions regarding the nature of the distribution, the variance and its behaviour over the range of response in replicated measurements under identical conditions. In addition intralaboratory reproducibility was assessed by the maintenance of minimum criteria, e.g. a detection limit below an a-priori chosen positive control or a dose-dependent standard curve (table 5.1.1, experiment 1b). With the data of this experiment an assessment of the limit of detection of the test can be done by calculating the smallest spike, which can be discriminated from the blank. Intralaboratory reproducibility was assessed by the maintenance of minimum criteria, e.g. a detection limit below an a-priori chosen positive control of a dose dependent standard curve.

The MM6/IL-6 method was transferred from the DL to two other laboratories (denoted as naive laboratory 1 [NL1] and naive laboratory 2 [NL2]). A large-scale dose response experiment was performed by all three laboratories. For this study 6 or 7 concentrations were tested in a dose response curve (typically 0; 0.125; 0.25; 0.5; 1; 2; 4 EU/ml, at least

8 replicates) and all laboratories had to meet the validity criteria as laid down in the MM6/IL-6 protocol before the studies with medicinal substances were conducted.

The **(intra- and interlaboratory) reproducibility** was assessed by testing 3 different medicinal substances, Gelafundin, Jonosteril and Haemate (described in table 3.3.2, section 3.3.). Test items and their spikes were appropriately blinded. Test items were tested, at a predefined dilution above the MVD, independently in 3 laboratories, 3 times each. Test items were tested after spiking with WHO-LPS at four different levels, the spikes were blinded and coded by QA ECVAM. In addition a negative control (saline) and positive control (0.5 EU/ml) in saline were included to establish assay validity. Although this part of the study was designed for assessment of reproducibility, a preliminary estimate of the accuracy could be derived from the data. Applying the PM to the results and evaluating the concordance in a two-by-two contingency table assessed accuracy.

To assess **accuracy** of the proposed test method 10 substances (listed in table 3.3.1), were spiked with five different concentrations of the WHO-LPS (one of which is negative). To permit the application of the chosen PM, each drug was diluted to its individual MVD, which has been calculated using the ELC to that drug (listed in table 3.3.1.). Each substance had their own specific set of endotoxin spike solution, ensuring that after spiking the undiluted drug, it contained 0.0; 0.25; 0.5; 0.5 and 1.0 EU/ml respectively when tested at its MVD. A detailed description of the sample preparation was supplied to the three independent laboratories (shown in table 5.1.2). To put more weight to the result of this part of the validation, the spikes were blinded and coded by QA ECVAM. The raw data of the MM6/IL6 assay are shown in paragraph 5.2 Accuracy was assessed by applying the PM to the results and evaluating the concordance in a two by two table. As intralaboratory reproducibility was (successfully) shown in previous experiments, only interlaboratory reproducibility was assessed in this phase.

Table 5.1.2: Sample preparation for the testing of 10 substances spiked with 5 different concentrations of WHO-LPS.

<table>
<tr><td colspan="3">Unblinded</td><td colspan="5">blinded</td></tr>
<tr><td colspan="3">dilution of drug up to MVD
⬇</td><td colspan="5">spiking of undiluted drug: 0.5 ml each</td></tr>
<tr><td>diluted drug</td><td>NPC</td><td>PPC</td><td rowspan="2">+ 23.3 µl of Spike 1</td><td rowspan="2">+ 23.3 µl of Spike 2</td><td rowspan="2">+ 23.3 µl of Spike 3</td><td rowspan="2">+ 23.3 µl of Spike 4</td><td rowspan="2">+ 23.3 µl of Spike 5</td></tr>
<tr><td rowspan="2">0.5 ml</td><td>+ 25 µl saline</td><td>+ 25 µl PPC-LPS-spike *</td></tr>
<tr><td></td><td>(final conc. = 50 pg/ml)</td><td colspan="5">dilution to MVD
⬇</td></tr>
<tr><td></td><td>test</td><td>test</td><td>test</td><td>test</td><td>test</td><td>test</td><td>test</td></tr>
</table>

* PPC-LPS-spike contains 1050 pg/ml = 21fold 50 pg/ml

NPC = Negative Product Control, **PPC** = Positive Product Control, **MVD** = Maximal Valid Dilution

5.2 Accuracy and reliability

Provide all data obtained to evaluate the accuracy and reliability of the proposed test method. This should include copies of original data from individual animals and/or individual samples, as well as derived data. The laboratory's summary judgment regarding the outcome of each test should be provided. The submission should include data (and explanations) from all studies, whether successful or not.

See figures 5.2.1, 5.2.2, 5.2.3, 5.2.4, 5.2.5 (A, B and C), 5.2.6 and 5.2.7 (A and B).

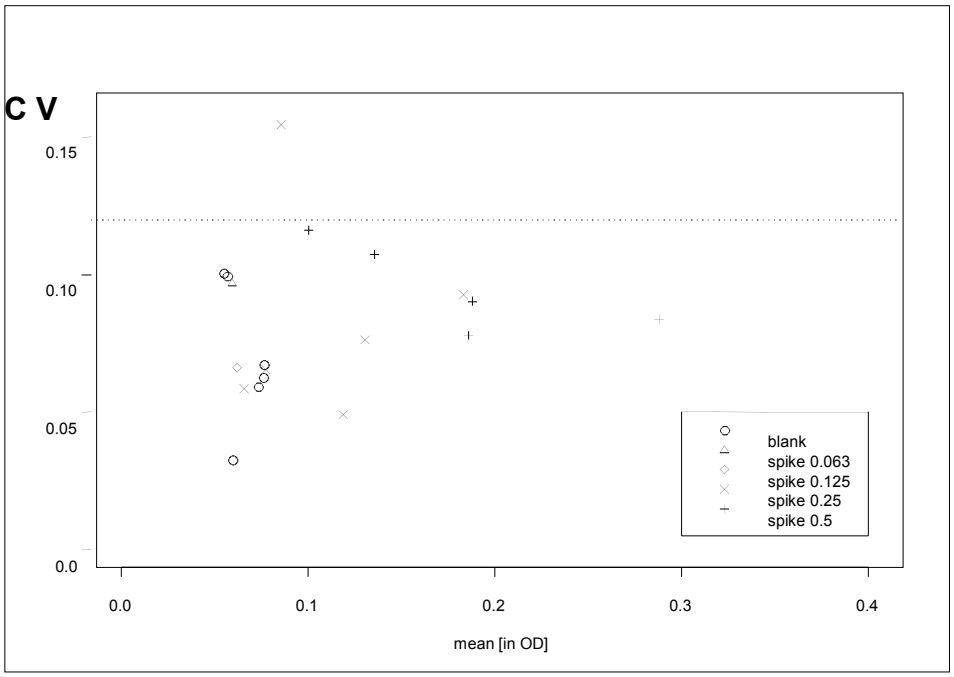


Figure. 5.2.1: Coefficient of variation (CV) of WHO-LPS spikes (4 replicates) relative to the mean OD (readout of the IL-6 ELISA).

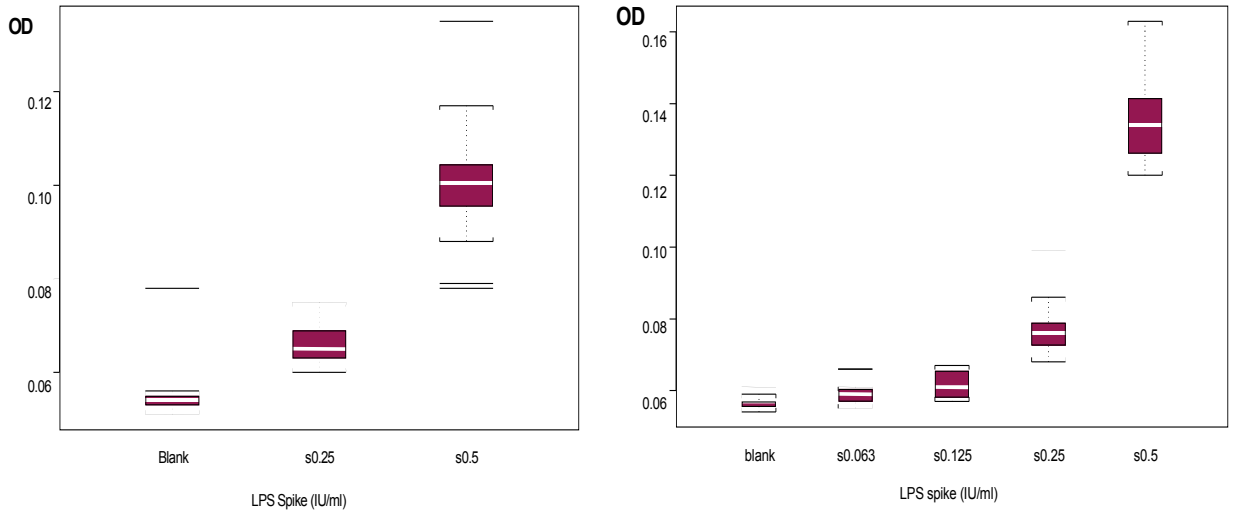
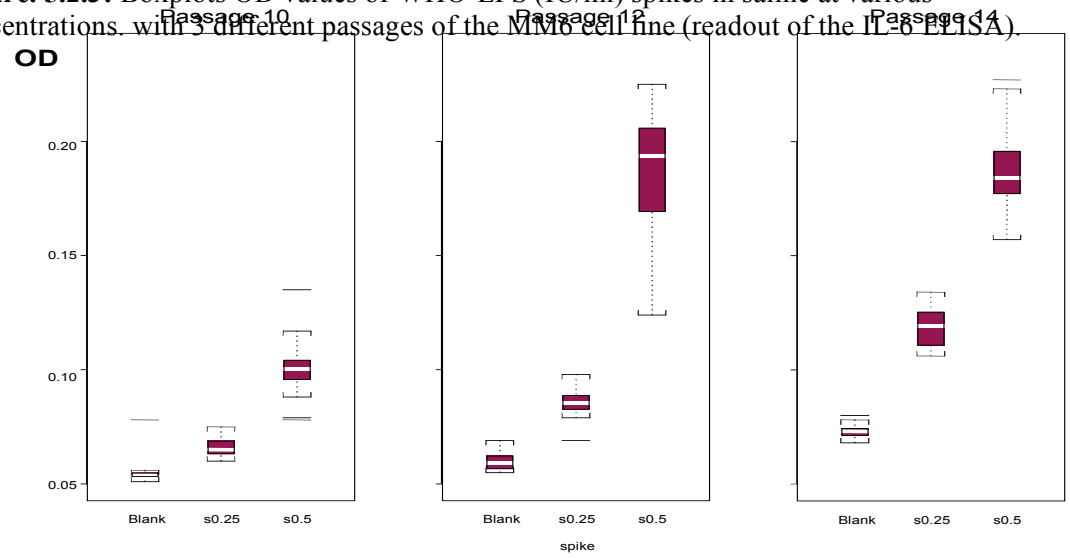


Figure. 5.2.2: Boxplots with OD values of 20- replicates (left) or 12 replicates (right) of WHO-LPS spikes in saline at various concentrations (readout of the IL-6 ELISA).

Figure. 5.2.3: Boxplots OD values of WHO-LPS (IU/ml) spikes in saline at various concentrations, with 3 different passages of the MM6 cell line (readout of the IL-6 ELISA).



BRD: MM6/IL-6 March, 2006

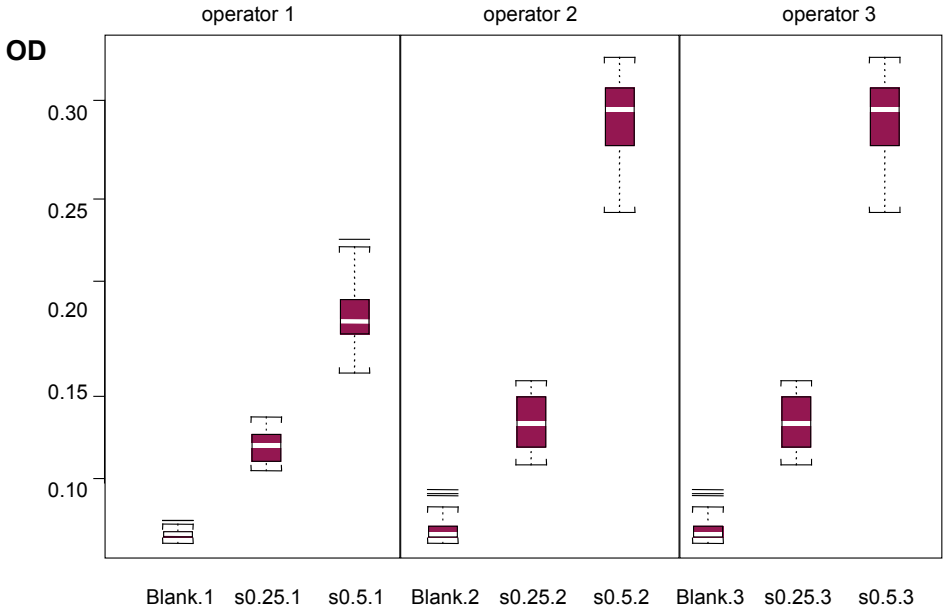


Figure. 5.2.4: Boxplots OD values of WHO-LPS (IU/ml) spikes in saline at various concentrations. with 3 different operators (readout of the IL-6 ELISA).

BRD: MM6/IL-6 March, 2006

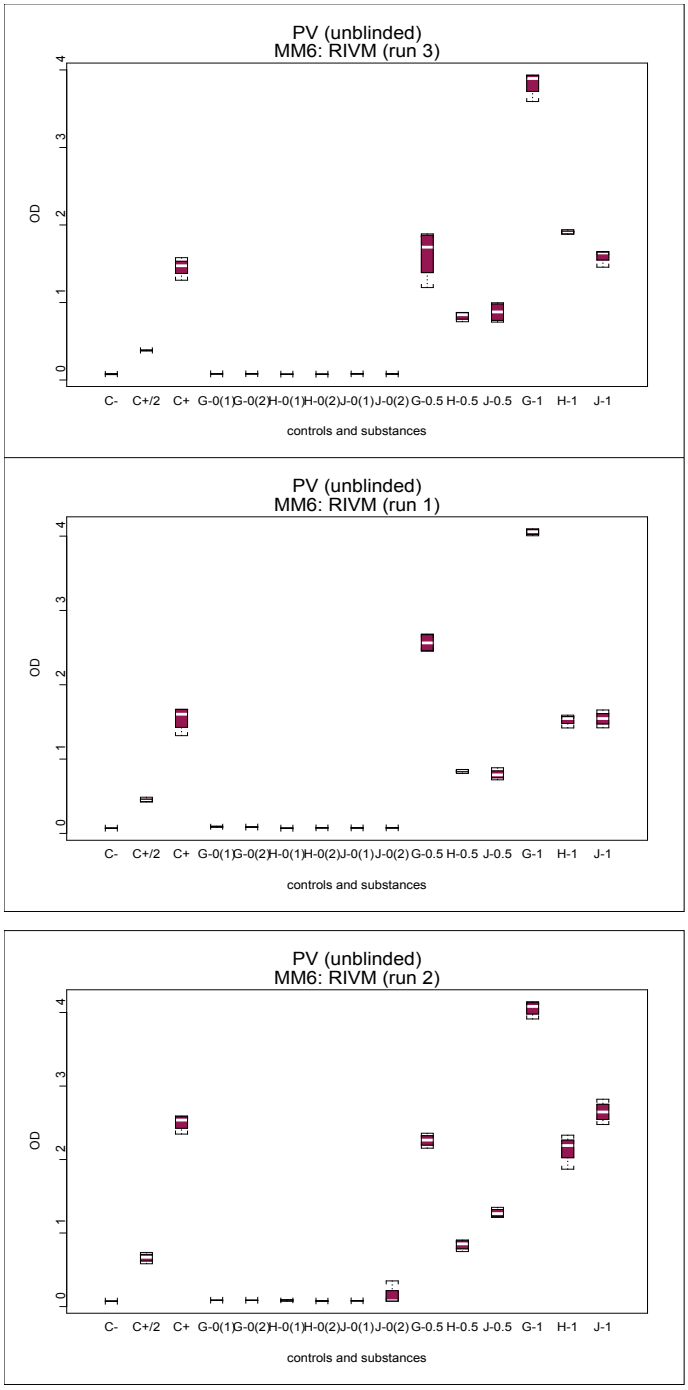
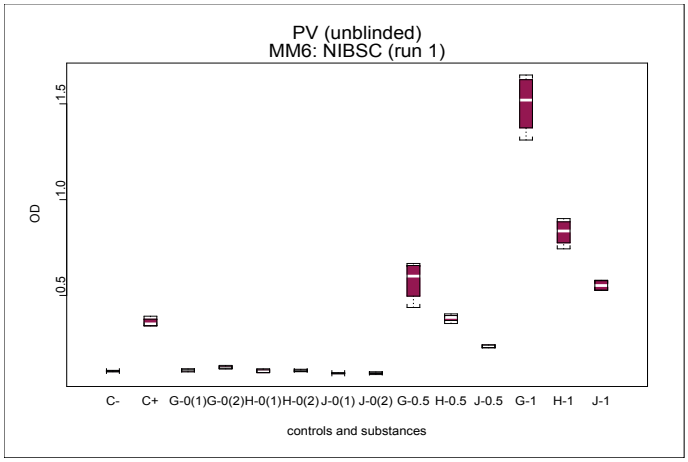
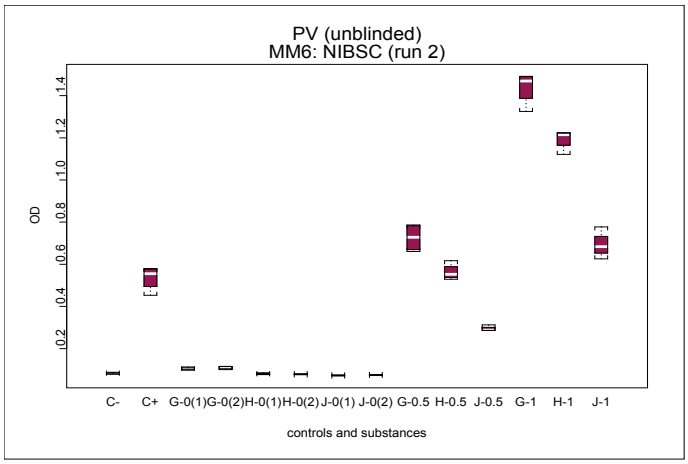


Figure. 5.2.5 A: Three different drugs (unblinded) spiked with 4 (blinded) concentrations of WHO-LPS (0.0, 0.0, 0.5 and 1.0 IU/ml, respectively). Experiment was run 3 time independently at three different laboratories. Here the results of the RIVM (readout of the IL-6 ELISA).
G = Gelafundin; J = Jonestreril; H = Heamate.
C- = negative control (saline); C+ = positive control (0.5 IU/ml in saline).





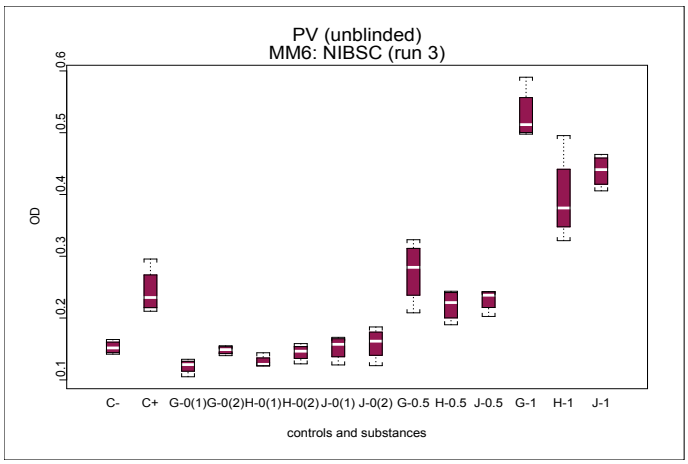
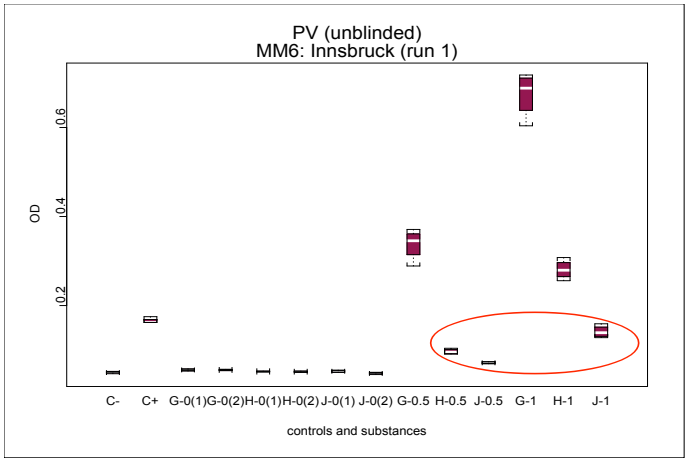
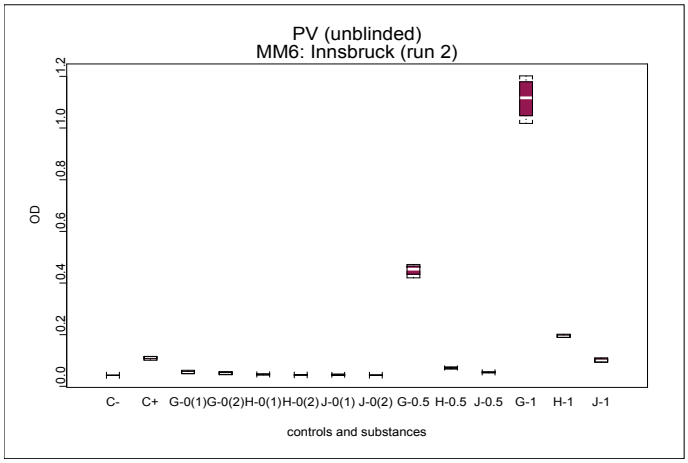


Figure. 5.2.5 B: Three different drugs (unblinded) spiked with 4 (blinded) concentrations of WHO-LPS (0.0, 0.0, 0.5 and 1.0 IU/ml, respectively). Experiment was run 3 time independently at three different laboratories. Here the results of the NIBSC (readout of the IL-6 ELISA). G = Gelafundin; J = Jonestreril; H = Heamate.
C- = negative control (saline); C+ = positive control (0.5 IU/ml in saline).





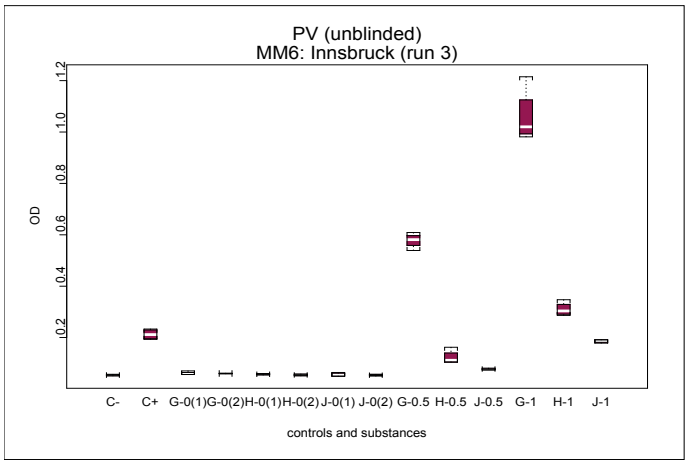


Figure. 5.2.5 C: Three different drugs (un-blinded) spiked with 4 (blinded) concentrations of WHO-LPS (0.0, 0.0, 0.5 and 1.0 IU/ml, respectively). Experiment was run 3 time independently at three different laboratories. Here the results of Innsbruck (readout of the IL-6 ELISA).
G = Gelafundin; J = Jonestreril; H = Heamate.
C- = negative control (saline); C+ = positive control (0.5 IU/ml in saline).

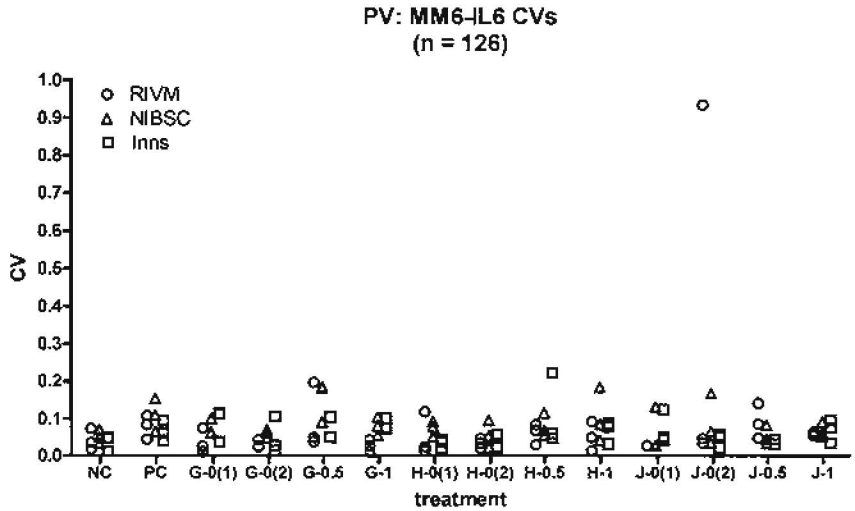


Figure. 5.2.6: Coefficient of variation (CV) of different WHO-LPS spikes (0.0, 0.0, 0.5 and 1.0 IU/ml, respectively).from the experiments as shown in fig. 5.2.5A-C.
G = Gelafundin; J = Jonestreril; H = Heamate.
NC = negative controle (saline); PC is positive conrole (0.5 IU/ml in saline).

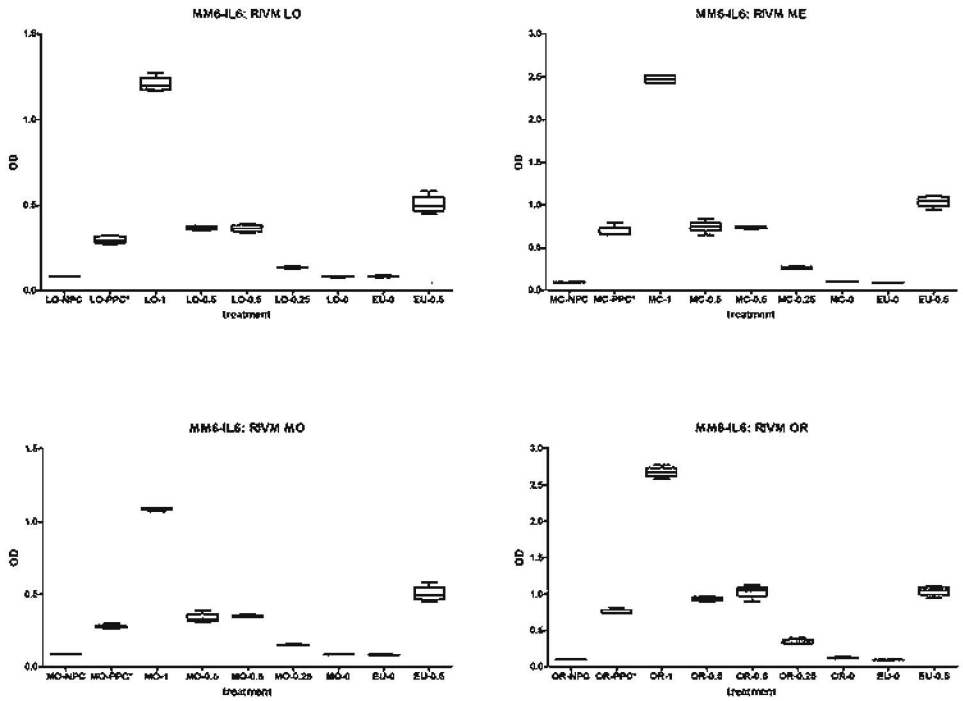


Figure. 5.2.7 A: Boxplots of the results of the validation study. Drugs LO, ME, MO and OR (see table 3.3.1) were spiked with WHO-LPS at 0.0, 0.0, 0.5 and 1.0 IU/ml, respectively.
NPC = negative product control; PPC = positive product control (0.5 IU/ml in drug)
EU-0 is negative control (saline); EU-0.5 = positive control (0.5 IU/ml in saline)

BRD: MM6/IL-6 March, 2006

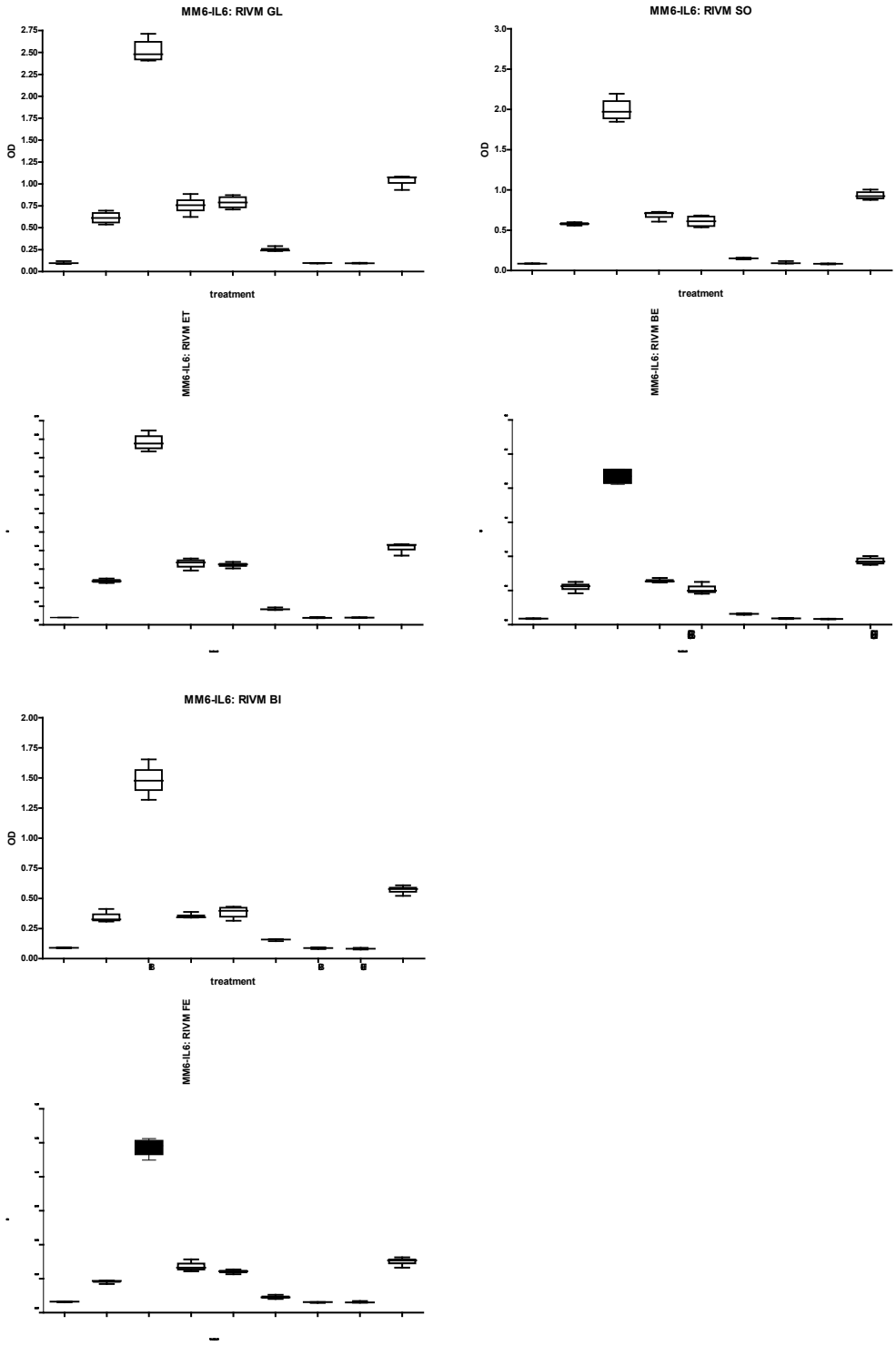


Figure. 5.2.7 B: Boxplots of the results of the validation study. Drugs GL, SO, ET, BE, BI and FE (see table 3.3.1) were spiked with WHO-LPS at 0.0, 0.0, 0.5 and 1.0 IU/ml, respectively. NPC = negative product control; PPC = positive product control (0.5 IU/ml in drug), EU-0 is negative control (saline); EU-0.5 = positive control (0.5 IU/ml in saline)

5.3 Statistics

Describe the statistical approach used to evaluate the data resulting from studies conducted with the proposed test method.

A generally applicable analytical procedure was employed. This procedure includes a universal PM as well as quality criteria. First a two-step procedure, consisting of a variance-criterion and an outlier-test was applied. For this, the Dixon's test (Barnett and Lewis, 1984), which is USP approved, was chosen with the significance level of α=0.01 and applied to identify and eliminate aberrant data.

Next, the negative and the respective positive control are compared to ensure a suitable limit of detection. For this, a one-sided t-test with a significance level of α=0.01 is applied to the ln-transformed data to ensure that the response to the positive control is significantly larger than that of the respective negative control.

Finally, the samples are classified as either negative or positive by the outcome of a one-sided version of the t-test, which is based on the assigned pyrogen threshold value. The final results will be given in 2 x 2 contingency tables (table 5.3.1). These tables allow for estimation of accuracy (sensitivity and specificity) and reproducibility of the proposed test method.

Table 5.3.1: 2x2 contingency table.

		pre-defined class ("truth")		Σ
		1	***0***	
Classification by test system and PM	***1***	a	b	$a+b = n_{.1}$
	0	c	d	$c+d = n_{.0}$
Σ		$a+c = n_{1.}$	$b+d = n_{0.}$	n

Accuracy:

The most important statistical tool to determine accuracy (specificity and sensitivity) is the so-called prediction model (PM) (Hothorn, 1995). In general, it is a statistical model, which classifies a given drug by an objective diagnostic or deciding rule. The objective of a dichotomous result requires a clear cut PM, which assigns a drug in one of the two classes "pyrogenic for humans" and "non-pyrogenic for humans". Since a threshold pyrogen value will be used, a one-sided test is appropriate for the task. Because the data are normalised by a ln-transformation, a t-test is chosen. Although the variances over the range of concentration converge by the transformation, the assumptions of equal variances do generally not hold true, because it depends on additional covariates.

Therefore, the one sided Welch-t-test (Snedecor and Cochran, 1989) is applied. Due to the safety aspect of the basic problem, the hypotheses of the test are

$$H_0 : \mu_{S_i j} > \mu_{S+} \quad vs \quad H_1 : \mu_{S_i j} < \mu_{S+},$$

where $\mu_{...}$ denotes the parameter of location of the respective ln-transformed distribution. This approach controls the probability of false positive outcomes directly by means of its significance level α, which is chosen as 0.01, because is assumes hazard, respectively pyrogenicity, of the tested drug in H_0, and assures safety, i.e. non-pyrogenicity. The test statistic is

$$T_{S_i j} = \frac{\bar{x}_{S+} - \bar{x}_{S_i j}}{\sqrt{\frac{s^2_{S+}}{n_{S+}} + \frac{s^2_{S_i j}}{n_{S_i j}}}}.$$

The PM is built by means of the outcome of the test. Let 0 denote safety and 1 denote hazard. The classification of S_i-j is then determined by

$$S_i j = 0, \text{ if } T_{S_i j} > t_{0.99; n_{S+} + n_{S_i j} - 2},$$

$$S_i j = 1, \text{ else,}$$

where $t_{0.99; n_{S+} + n_{S_i j} - 2}$ the 0.99-quantile of the t-distribution with $n_{S+} + n_{S_i j} - 2$ degrees of freedom. The number of replicates for every control and sample, i.e. $n_{...}$, was harmonised to be four. Due to the possibility of removing one observation by the outlier test, the number of replicates could be reduced to three. The classification of a version of a drug is regarded as an independent decision. Therefore, the niveau α is local.

Finally, the classifications of the drugs will be summarised in 2x2 contingency table (table 3). From these tables, estimates of the sensitivity (S_E), i.e. the probability of correctly classified positive drugs and specificity (S_P), i.e. the probability of correctly classified negative drugs, will be obtained by the respective proportions. Where

$$S_E = a / (a + c) * 100\%$$
and
$$S_P = d / (b + d) * 100\%.$$

Furthermore, these estimates will be accompanied by confidence intervals, which will be calculated by the Pearson-Clopper method (Clopper & Pearson, 1934). For example, let $\hat{p}_{SE}$ denote the proportion, namely the sensitivity, under investigation. Then the confidence interval to a niveau α is calculated as

$$\left[p_{SE}^{L} = \frac{\mathrm{a}F_{2\mathrm{a};2(n_{1\cdot}-\mathrm{a}+1);\frac{\alpha}{2}}}{n_{1\cdot}-\mathrm{a}+1+\mathrm{a}F_{2\mathrm{a};2(n_{1\cdot}-\mathrm{a}+1);\frac{\alpha}{2}}}; p_{SE}^{U} = \frac{(\mathrm{a}+1)F_{2(\mathrm{a}+1);2(n_{1\cdot}-\mathrm{a});1-\frac{\alpha}{2}}}{n_{1\cdot}-\mathrm{a}+(\mathrm{a}+1)F_{2(\mathrm{a}+1);2(n_{1\cdot}-\mathrm{a});1-\frac{\alpha}{2}}} \right],$$

where $F_{\ldots}$ denotes the respective quantile of the F-distribution and $n_{1\cdot}$ is the sample size of the positive drugs and a the number of correctly classified drugs.

By contaminating the drugs artificially and by defining a threshold value, which is assumed to be appropriate, the class of a drug is determined beforehand. The versions of drugs, which are effectively contaminated, but below the threshold dose, are considered to be negative, respectively safe, because their contamination is not crucial for humans in terms of ELC.

Reproducibility:

The analysis of the intra- and interlaboratory reproducibility was assessed from the three identical and independent runs conducted in each of 3 laboratories. The comparison of the three runs was carried out blindly such that the testing facility did not know the true classification of the sample, either pyrogenic or non-pyrogenic. By this procedure only the randomness and reproducibility of the methods was assessed and not systematic errors, which may have arisen from other sources, e.g. logistics or preparation of the samples. This means that if a sample was classified in all three runs the result is reproducible regardless of the classification of the sample. Therefore, a measure of similarity, i.e. complete simple matching with equal weights, was preferred to the coefficient of correlation for 2x2 contingency tables.

The study was designed as follows: each laboratory had to conduct three independent runs with the same 12 samples (3 test items with 4 blinded spikes each) and two controls, i.e. saline as a negative control (C-) and a 0.5 EU/ml LPS-spike in saline as a positive controls (C+). The samples were derived from the three substances Gelafundine, Haemate and Jonosteril. Per run, each substance was blindly spiked twice with saline, once with 0.5 EU/ml LPS and once with 1 EU/ml LPS, which resulted in a balanced design with regard to positive and negative samples, i.e. samples expected to be pyrogenic and non-pyrogenic, respectively.

The three independent runs per testing facility provide the information on which the assessment of the intralaboratory reproducibility is based. The combined results of the three runs per testing facility were used to determine interlaboratory reproducibility. The correlation of the prediction (in terms if the Bravais-Pearson coefficient of correlation) between all runs is calculated, independent of whether that classification is true or false. A BP-correlation of 1 is calculated, if two runs gave exactly the same predictions for the twelve substances. If one run gives adverse classifications for all substances than the other, the correlation is –1. As these calculations do not need information of the true status of a sample, they were carried out blinded.

5.4 Tabulated results

Provide a summary, in graphic or tabular form, of the results.
See tables 4.4.1 and 5.4.2.

Table 5.4.1: Results of testing 3 substances 3 times by 3 laboratories. Classifications after applying the prediction model (compare to fig. 5.2.5)

Sample	DL (RIVM)			NL 1 (NIBSC)			NL 2 (Innsbruck)		
	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3
G-0 (1)	0	0	0	0	0	0	0	0	0
G-0 (2)	0	0	0	0	0	0	0	0	0
H-0 (1)	0	0	0	0	0	0	0	0	0
H-0 (2)	0	0	0	0	0	0	0	0	0
J-0 (1)	0	0	0	0	0	0	0	0	0
J-0 (2)	0	0	0	0	0	0	0	0	0
G - 0.5	1	1	1	1	1	1	1	1	1
H - 0.5	0	0	0	0	0	0	1	1	1
J - 0.5	0	0	0	0	0	0	0	0	1
G - 1	1	1	1	1	1	1	1	1	1
H - 1	1	1	1	1	1	1	1	1	1
J - 1	1	1	1	0	1	1	1	1	1

"0"denotes "non-pyrogenic"; "1" denotes "pyrogenic".

Table 5.4.2: Results of the validation study of 10 drugs, spiked with WHO-LPS at 0.0, 0.25, 0.5, 0.5 and 1.0 IU/ml, respectively and tested in 3 different laboratories. Samples and spikes were blinded. Classifications after applying the PM (compare to fig. 5.2.7).

drug (code)	spike		results		
	EU/ml	"truth"	rivm	inns	nibsc
Beloc (BE)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**1**
	0.50	**1**	**1**	**1**	**1**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
Binotal (BI)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**0**
	0.50	**1**	**1**	**1**	**1**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
Ethanol 13% (ET)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**1**	**0**
	0.50	**1**	**1**	**1**	**1**
	0.50	**1**	**1**	**1**	**1**
	1.00	**1**	**1**	**1**	**1**
Fenistil (FE)	**0.00**	**0**	**0**	**0**	**0**
	0.25	**0**	**0**	**0**	**0**
	0.50	**1**	**1**	**0**	**1**

BRD: MM6/IL-6 March, 2006

drug (code)	spike EU/ml	"truth"	results rivm	inns	nibsc
	0.50	1	1	0	1
	1.00	1	1	1	1
Glucose 5% (GL)	0.00	0	0	0	0
	0.25	0	0	1	0
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
"Drug A" 0.9% NaCl (LO)	0.00	0	0	0	0
	0.25	0	0	0	0
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
MCP (ME)	0.00	0	0	0	0
	0.25	0	0	0	0
	0.50	1	1	1	0
	0.50	1	1	1	0
	1.00	1	1	1	1
"Drug B" 0.9% NaCl (MO)	0.00	0	0	0	0
	0.25	0	0	1	1
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1
Orasthin (OR)	0.00	0	0	0	0
	0.25	0	0	0	1
	0.50	1	1	1	1
	0.50	1	1	1	NA
	1.00	1	1	1	1
Sostril (SO)	0.00	0	0	0	0
	0.25	0	0	0	NA
	0.50	1	1	1	1
	0.50	1	1	1	1
	1.00	1	1	1	1

"0"denotes "non-pyrogenic"; "1" denotes "pyrogenic".

5.5 Coding of data

For each set of data, indicate whether coded chemicals were tested, whether experiments were conducted without knowledge of the chemicals being tested, and the extent to which experiments followed GLP guidelines.
Blinding of drugs and/or spikes is indicated with the data.

5.6 Circumstances

Indicate the "lot-to-lot" consistency of the test substances, the time frame of the various studies, and the laboratory in which the study or studies were conducted. A coded designation for each laboratory is acceptable.
Per part of the study, all samples are derived from one (clinical) lot.

5.7 Other data available

Indicate the availability of any data not submitted for external audit, if requested.
All relevant data were submitted with the present BRD.

6 Test Method Accuracy

6.1 Accuracy

Describe the accuracy (e.g., concordance, sensitivity, specificity, positive and negative predictivity, false positive and negative rates) of the proposed test method compared with the reference test method. Explain how discordant results in the same or multiple laboratories from the proposed test were considered when calculating accuracy.
Test method accuracy was assessed in two large scale experiments performed with the drugs outlined in table 3.3.1 and table 3.3.2 in section 3 respectively. As described before one experiment was performed in an early stage of the study with 3 different drugs, tested 3 times and the other final experiment all drugs were tested once in the three participating laboratories. From the first experiment a preliminary estimate of sensitivity and specificity can be figure out, whereas the second is regarded as the established accuracy for the MM6/IL6 assay.

6.1.1 Preliminary estimate of the accuracy of the MM6/IL6 assay. In an early stage of the study a different concept for interference testing was used. The developing laboratories (DLs) determined for each drug (outlined in table 3.3.2, section 3.3) the smallest dilution within the MVD that showed no interference or an acceptable degree of interference with the spike recovery. In general the lowest dilution of the sample allowing for a 50-200% spike recovery was chosen. In addition, the positive control (PC) set at 0.5 EU/ml saline was used as the classification threshold. The laboratory procedure as described in the Protocol was maintained throughout the study. Although it was realized there were some drawbacks to the concept for interference testing and applying the PC as a threshold, this small scale study allows for a preliminary estimate of the accuracy of the MM6/IL6 method.
It has to be noted that this part of the study was designed to provide an estimate of the intra and inter laboratory reproducibility. Therefore it will also be discussed in detail in section 7 (Test Method Reliability).

According to the PM applied during an early phase of the study the outcome (positive/negative) is related to the positive control (PC=0.5 EU/ml). If absorbance of sample > absorbance of PC, then the sample is classified as positive. If absorbance of sample < PC, then the sample is classified as negative. While performing the experiments during this phase it was realized that there is a flaw between interference testing and the PM. After the interference testing, the lowest dilution of the sample allowing for a 50-200% spike recovery was chosen. It is obvious that even with a flawless repeatability of the assay; a spike recovery between 50%-100% would be classified as negative according to the preliminary PM. In addition, due to unforeseen problems with the preparations of the spike, the recovery of the spikes was far below 100%. (This is outside the scope of the study and will not be discussed). As a consequence of the employed preliminary setup of the study the sensitivity will be underestimated, and the specificity will be overestimated.

In short, three test substances were spiked with WHO-LPS at four levels (0, 0, 0.5 and 1.0 EU/ml respectively), which resulted in 12 samples with a balanced design with regard to positive and negative samples (i.e. samples expected to be pyrogenic and non-pyrogenic respectively. These 12 sample were three times tested in three laboratories. In total there were 108 classifications from 12 samples in 3 runs and in 3 laboratories (3x3x12=108). Results are described in detail in section 7. A 2x2 contingency table was constructed (table 6.1.1), from which the estimates of sensitivity and specificity can easily be derived.

Table 6.1.1: 2x2 contingency table. The prediction model applied to a preliminary study.

		True status of samples		**Total**
		+	-	
PM	+	39	0	39
	-	15	54	69
	Total	54	54	108

The specifications of specificity and sensitivity described in section 5.3 were applied to these results and the specificity (Sp) of the MM6/IL6 assay is 100% (54/(54+0)*100%), 95% confidence interval [0.934;1]. The sensitivity (Se) equals 72% (39/(39+15) *100%), 95% confidence interval [0.602; 0.859]. As outlined previously the specificity is overestimated and the sensitivity is underestimated as a result of the design of this part of the study.

6.1.2 Test method accuracy of the proposed MM6/IL6 assay. To assess accuracy of the proposed MM6/IL6 method, 10 substances (listed in table 3.1.1, section 3) were spiked with five different concentrations of the WHO-LPS (one of which is negative). Thus, in total, 50 samples have been tested in each laboratory.
To permit the application of the chosen PM, each drug was diluted to its individual MVD, which has been calculated using the ELC to that drug (listed in section 3). Lesser dilutions were tested by the DL, and showed no interference. Therefore interference was not expected at the individual MVD. Each substance had their own specific set of endotoxin spike solution, ensuring that after spiking the undiluted drug, it contained 0.0; 0.25; 0.5; 0.5 and 1.0 EU/ml respectively when tested at its MVD. A detailed description of the sample preparation was supplied to the three independent laboratories (shown in table 5.1.1 for convenience). To put more weight to the result of this part of the validation, the spikes were blinded and coded by QA ECVAM. The raw data and the graphical presentation of these raw data are shown in the section 5 (table 5.4.2 and figure 5.2.7). Accuracy was assessed by applying the PM to the results (summarized in table 5.3.2) and evaluating the concordance in this section in a two by two contingency table (table 6.1.2). As described above 10 substances, spiked with 5 different WHO-LPS concentrations were tested in three laboratories and consequently a maximum of 150 data were available for analysis.

As intralaboratory reproducibility was (successfully) shown in previous experiments (analysed in section 7), only one run performed in each laboratory was considered sufficient.

Of the 150 available data, only two sets of 4 replicates did not comply with the quality criteria as defined in the PROTOCOL (CV >0.25). The specificity and sensitivity of the MM6/IL6 assay could be estimated as described in section 5.3.

Table 6.1.2:
2x2 contingency table. Prediction model applied to the MM6/IL6 test result of 10 different substances assessed in three different laboratories.

		True status of samples +	-	Total
PM	+	85	6	91
	-	4	53	57
	Total	89	59	148

The specificity of the MM6/IL6 assay is 89.8% (53/(53+6)*100%), 95% confidence interval [0.792;0.962]. The sensitivity equals 95.5% (85/(85+4) *100%), 95% confidence interval [0.889;0.998] (See table 6.1.3). The specificity varied from 84% up to 100% within the three laboratories, and the sensitivity varied from 93% up to 100%.

Table 6.1.3: Specificity and sensitivity of the MM6/IL6 assay

	N total	N correctly identified	proportion	95% CI lower limit	95% CI upper limit
Specificity (Sp)	59	53	89.8%	79.2%	96.2%
Sensitivity (Se)	89	85	95.5%	88.9%	99.8%

6.2 Concordancy to *in vivo* reference method

Discuss results that are discordant with results from the in vivo reference method.
Not applicable.

6.3 Comparison with reference methods

Discuss the accuracy of the proposed test method compared to data or recognized toxicity from the species of interest (e.g., humans for human health-related toxicity testing), where such data or toxicity classification are available. This is essential when the method is measuring or predicting an endpoint for which there is no preexisting method. In instances where the proposed test method was discordant from the in vivo reference test method, describe the frequency of correct predictions of each test method compared to recognized toxicity information from the species of interest.
Not applicable.

6.4 Strength and limitations

State the strengths and limitations of the proposed test method, including those applicable to specific chemical classes or physical-chemical properties.

It appears the proposed test is applicable to most classes of medicinal products, at least those that are non- or low-toxic to cells *in vitro*. In addition, the test may be employed to assess pyrogenicity of various medical devices, such as (biological) bovine collagen bone implants.

6.5 Data interpretation

Describe the salient issues of data interpretation, including why specific parameters were selected for inclusion.
No issues.

6.6 Comparison to other methods

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the results obtained with both test methods should be compared with each other and with the in vivo reference test method and/or toxicity information from the species of interest.
Not applicable.

7 Test Method Reliability (Repeatability/Reproducibility)

7.1 Selection of substances

Discuss the selection rationale for the substances used to evaluate the reliability (intralaboratory repeatability and intra- and interlaboratory reproducibility) of the proposed test method as well as the extent to which the chosen set of substances represents the range of possible test outcomes.
The rationale for the selection of the substances is described in section 3.3. In short: for the present studies endotoxin (WHO-LPS) was selected as the model pyrogen, since it is well defined biological standard and readily available. Selected test substances were medicinal products available on the market. These batches are released by the manufacturers and comply with the Marketing Authorisation file and European Pharmacopoeia. Therefore these batches are considered to contain no detectable pyrogens. To test the method reliability the medical products were spiked with endotoxin.

7.2 Results

Provide analyses and conclusions reached regarding the repeatability and reproducibility of the proposed test method. Acceptable methods of analyses might include those described in ASTM E691-92 (13) or by coefficient of variation analysis.
The following definitions according to the ICCVAM guideline were applied: Reliability is assessed by calculating intra- and interlaboratory reproducibility: replicate results at different times (persons). Intralaboratory repeatability: no variation, identical conditions within a given period.

In an early phase of the study, the intralaboratory repeatability and reproducibility of the test method was assessed in a series of experiments conducted in the DL. Series of blanks (saline, 0 EU/ml) and spikes of WHO-LPS in saline were tested. These experiments are summarized in table 5.1.1. The data were used to answer questions regarding the nature of the distribution, the variance and its behavior over the range of response in replicated measurements under identical conditions. In addition reliability of the test method was assessed by the maintenance of minimum criteria, e.g. a detection limit below an a-priori chosen positive control or a dose-dependent standard curve With the data of this experiment an assessment of the limit of detection of the test can be done by calculating the smallest spike, which can be discriminated from the blank.

The second group of experiments was meant to analyze the variation in detail. For this purpose the major sources of variation were assessed separately i.e.: Passage-number of cell-line (fig. 5.2.3) and operator (fig. 5.2.4). As some experiments were combined, a total of 360 data were collected and analyzed.

First the shape of the distribution at a spike was assessed (not shown). Most of the data showed a right-skewed shape which suggests that the data should be analyzed by a parametric approach via the lognormal-distribution.
Based on the experience that there is a monotone increasing relationship between the mean-responses and the variation (empirical variance or standard deviation), the analysis

focuses on the coefficient of variation (CV). The CV should be distributed symmetric around a constant factor, if the mean-variance relationship is linear. A plot of all CVs against their corresponding means is shown in figure 5.2.1. From the figure it is clear that the coefficient of variation for a set of 4 replicates of one spike concentration is usually below 8% (CV < 0.12), which is considered very small. As only WHO-LPS was examined up to this point, it was envisaged that the CV would increase with other substances being tested. For CV criteria applied as a validity criteria of the MM6/IL6 assays was arbitrarily set at CV<0.25.

After a log-transformation of all data obtained with all four experiments (1a, 1b, 2a and 2 b), a parametric test for detection of outliers was applied. At this point the Grubbs-test was chosen. Altogether there were 3 outliers, which is 1% of the data analysed. In addition, the raw data (plate-readouts) showed no obvious edge effects or trends.

The results of test 1a and 1b (fig. 5.2.2) show that the 0.25 EU/ml of spike can be discriminated statistically from the blank and the highest spike (0.5 EU/ml) can be detected easily.

Test 2a was designed to assess the effect of the passage of the cell-line. Passage 10, 12 en 14 were tested (fig. 5.2.3). The general behavior of the assay with different passages is always the same, however the responses for the 0.5 EU/ml-spike differ in height (ELISA read out).

Experiment 2b (fig. 5.2.4) revealed an effect of the operator, but still the limit of detection is about 0.25 EU/ml and 0.5 EU/ml can be easily discriminated from the blanks.

In conclusion: The intralaboratory repeatability is satisfactory as the MM6/IL-6 assay shows a low variability while testing blanks and spikes with WHO-LPS. The experiments revealed an effect for the covariates passage and operator, but the limit of detection is about 0.25 EU/ml for all experiments, thus 0.5 EU/ml is always detectable. Therefore the intralaboratory repeatability is considered satisfactory. The 1% percentage outliers, as determined by the Grubbs test is very acceptable. The validity criteria of the MM6/IL-6 assay as recorded in the protocol, are based on these experiments (i.e. CV< 0.25, lower limit of detection 0.25 EU/ml).

Intra- and interlaboratory reproducibility.
After transfer of the MM6/IL-6 assays to two other laboratories, a dose response experiments was performed by all three laboratories. For this study 6 or 7 concentrations were tested in a dose response curve (typically 0, 0125, 0.25, 0.5, 1, 2, 4 EU/ml, at least 8 replicates). A participating laboratory qualified for taking part in next part of the study by producing a dose response curve, with a limit of detection of at least 0.25 EU/ml and a CV < 0.25. (Data not shown).

The intra- and interlaboratory reproducibility was assessed by testing 3 different medicinal substances, Gelafundin, Jonosteril and Haemate (described in table 3.3.2, section 3.3.). Test substances and their spikes were appropriately blinded. Test substances

were tested, at a predefined dilution above the MVD, independently in 3 different laboratories, 3 times each. The three test substances were spiked with WHO-LPS at four levels (0, 0, 0.5 and 1.0 EU/ml respectively), which resulted in 12 samples with a balanced design with regard to positive and negative samples (i.e. samples expected to be pyrogenic and non-pyrogenic respectively. In addition a negative control (saline) and positive control (0.5 EU/ml in saline) were included to establish assay validity. To avoid interference, the DL performed interference testing in terms of the BET, i.e. 50-200% spike recovery, and decided on the dilution of the test substances. Dilutions chosen for Gelafundine, Haemate, Jonosteril were 1:2, 1:20 and 1:2 respectively. The data derived by the RIVM are taken as an example of the three laboratories. The raw data and a graphical presentation of the absorbance values are shown in section 5 (raw data exp.5 and fig. 5.2.5).

From the experiment with LPS-WHO only it was already concluded that the coefficient of variation was low for the MM6/IL-6 assay (CV < 0.12). It was envisaged that the CV was likely to be higher when testing different substances (different matrices) and was assessed for the current set of data. A plot of all CVs for all sets of 4 replicates of a drug with a spike is shown in fig. 5.2.6. From the figure it is clear that the coefficient of variation for a set of 4 replicates of one spike concentration is usually below 0.22, which is considered more than acceptable for a biological assay. There was only one outlier (CV>0.95) noticed for the J-0(2), which is probably due to a pipetting error. For the remainder of the studies the CV criteria applied as a validity criteria of the MM6/IL-6 assays was arbitrarily set at CV<0.25.

The analysis of the intralaboratory reproducibility was assessed from the three identical and independent runs conducted in each laboratory. The comparison of the three runs was carried out blindly such that the laboratory did not know the true classification of the sample (either pyrogenic or non-pyrogenic). By this procedure only the randomness and reproducibility of the methods was assessed and not systematic errors, which may have arisen from other sources, e.g. logistics or preparation of the samples. This means that if a sample was misclassified in all three runs the result is 100% intralaboratory reproducible (regardless of the misclassification of the sample).

According to the preliminary PM applied during this phase of the study the outcome (positive/negative) was related to the positive control (PC=0.5 EU/ml). If absorbance of sample > absorbance of PC, then sample is classified as being positive. If absorbance of sample < PC, sample is classified as negative. (positive/pyrogenic = 1, negative/non-pyrogenic = 0).
During this phase it was realized that there is a flaw between interference testing and the PM. After the interference testing, the lowest dilution of the sample allowing 50-200% spike recovery was chosen. It is obvious that even with a flawless repeatability of the assay, a spike recovery between 50%-100% would be classified as negative according to the preliminary PM.

From the result summarized in table 5.4.1, the intralaboratory reproducibility can be calculated for the separate laboratories. For these calculations there is no need for

information of the true status of the sample. A minimum criterion for the establishment of an assay is that experiment carried out with the same samples should result in a high concordance of classifications. From table 7.2.1 it can be read that the intralaboratory reproducibility very good (94.4%-100%) for all three participating laboratories.

Table 7.2.1 : Intralaboratory reproducibility, assessed by correlation between different runs. Result of testing 3 substances 3 times by 3 laboratories.

	DL	NL1	NL2
Run 1 - Run 2	100% (12/12)	92% (11/12)	100% (12/12)
Run 1 - Run 3	100% (12/12)	92% (11/12)	92% (11/12)
Run 2 - Run 3	100% (12/12)	100% (12/12)	92% (11/12)
Mean	100%	94.4%	94.4%
Proportion showing the same result in 3 runs	100%	92%	92%

The interlaboratory reproducibility of the MM6/IL6 method was assessed in a similar manner to the intralaboratory reproducibility. A summarizing method to combine the three runs per laboratory is considered not appropriate, because it would mask misclassification. Therefore each run of one laboratory was compared with all runs of another laboratory. The results summarized in table 7.2.2, show that there is a good interlaboratory reproducibility of at least 86%.

Table 7.2.2: Interlaboratory reproducibility. assessed by interlaboratory correlations. **Result of testing 3 substances 3 times by 3 laboratories.**

Laboratories	Interlab Reproducibility	Number of equal predictions
DL – NL1	97%	105 / 108
DL – NL2	89%	96 / 108
NL1 – NL 2	86%	93 / 108
Mean	90%	
Same result in all laboratories	86%	93 / 108

Also from the result of the large scale study (testing 10 substances spiked with 5 separate spikes), the *interlaboratory* reproducibility can be estimated (table 7.2.3). All the samples were correctly identified by one of the laboratories (DL). The reproducibility varied from 83.3% to 90% between two laboratories. The same results was found by all three laboratories for 39 out of 48 samples (81.3%).

Table 7.2.3: Interlaboratory reproducibility: Assessed by testing of 10 substances, spiked 5 times. One run of 50 samples by three different laboratories.

	Interlab reproducibility	Number of equal predictions
DL - NL1	90%	45 / 50
DL - NL2	90%	43 / 48
NL1 – NL2	83%	40 / 48
Mean	88%	
same result in all laboratories	81%	39 /48

Conclusion: It is shown that the intralaboratory reproducibility, assessed by correlation between different runs varies from 92% to 100% between the three participating laboratories. The interlaboratory reproducibility between two laboratories varied from 86% to 97% in one experiment and from 83% to 90% in the other. All three participating laboratories predicted the same in respectively 86% and 81% of the measurements. It has to be noted that part of the samples was 0.5 EU/ml and close to the arbitrary point of the MM6/IL-6 assay.

7.3 Historical data

Summarize historical positive and negative control data, including number of experiments, measures of central tendency, and variability.
Not applicable.

7.4 Comparison to other methods

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the reliability of the two test methods should be compared and any differences discussed.
Not applicable.

8 Test Method Data Quality

8.1 Conformity

State the extent of adherence to national and international GLP guidelines (7-12) for all submitted data, including that for the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method. Information regarding the use of coded chemicals and coded testing should be included.
The studies were done in accordance to the guidelines for GLP. Written protocols and approved standard operating procedures were followed during the entire course of the study. Deviations were recorded and, where appropriate, approved in amendments. All data are stored and archived. As mentioned, samples were appropriately blinded.

8.2 Audits

Summarize the results of any data quality audits, if conducted.
No audits were done.

8.3 Deviations

Discuss the impact of deviations from GLP guidelines or any noncompliance detected in the data quality audits.
Not applicable.

8.4 Raw data

Address the availability of laboratory notebooks or other records for an independent audit. Unpublished data should be supported by laboratory notebooks.
All records are stored and archived by the contributing laboratories and available for inspection.

BRD: MM6/IL-6 March, 2006

9 Other Scientific Reports and Reviews

9.1 Summary

Summarize all available and relevant data from other published or unpublished studies conducted using the proposed test method.

An *in vitro* monocyte activation test that detected pro-inflammatory and pyrogenic contaminants, was first applied some 15 years ago (Poole et al., 1988). A number of variants of the original test system have since been described, although the underlying principle of each variant remains the same. The test preparation is cultured with monocytes, either as peripheral blood mononuclear cells, PBMC, diluted whole blood or cells of a monocytoid cell line such as MM6. Contaminants in the test article activate CD14/TLR receptors which stimulates the release of an endogenous pyrogenic cytokine from the monocytes (Poole and Gaines Das, 2001).

Early studies report on opimization of the test method, e.g. improving the lower limit of detection, incubation times and cytokine readout, using model pyrogens such as LPS or endotoxin. Limited information is available on the actual testing of medicinal products.

Most interestingly, Taktak et al (1991) described several batches of a medicinal product (serum albumin) that caused adverse (pyrogenic) reactions in recipients. These lots were not detected by either BET or rabbit test but only by the *in vitro* monocytoid cell test.

In a study using whole blood and monocytoid cell lines as the sources of monocytoid cells (Nakagawa et al., 2002) it was reported that the structurally diverse pyrogens endotoxin, peptidoglycan, Staphylococcus aureus Cowan 1 and poly(I.C) all stimulated the release of cytokines.

The cytokine readout included tumour necrosis factor (TNF), interleukin-1 (IL-1), and IL-6 (reviewed by Poole and Gaines Das, 2001 and Poole et al., 2003). Other cytokines, e.g. IL-8, are also produced in large quantities in response to pyrogenic contaminants but their roles in fever are less well studied. The preferred readout is usually IL-6 because IL-6, unlike IL-1 and TNF, is secreted entirely into the cell-conditioned medium in large quantities, thereby permitting its complete estimation (Poole et al, 1988; Poole et al., 1989, Taktak et al., 1991).

Also, certain pro-inflammatory bacterial components stimulate the production of IL-6 but not TNF and IL-1 (Reddi et al., 1996), and IL-6 induction via Toll-like (pyrogen) receptors rapidly follows the recognition of microbial products (Pasare and Medzhitov, 2003).

9.2 Discussion

Comment on and compare the conclusions published in independent peer-reviewed reports or other independent scientific reviews of the proposed test method. The conclusions of such scientific reports and reviews should be compared to the conclusions reached in this submission. Any ongoing evaluations of the proposed test method should be described.

This is the first time such an extensive study for specificity and accuracy using actual medicinal products is carried out with the present method. Hence, there are no comparing reports in independent peer-reviewed journals available.

9.3 Results of similar validated method

In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the results of studies conducted with the validated test method subsequent to the ICCVAM evaluation should be included and any impact on the reliability and accuracy of the proposed test method should be discussed.

As mentioned, *in vitro* monocytoid activation test methods for the detection of pyrogenic contaminants are being developed over the course of the past two decades. A number of variants have been described, although the underlying principle of each variant remains the same. The test preparation is cultured with monocytoid cells, either as peripheral blood mononuclear cells, PBMC, (diluted) whole blood or cells of a monocytoid cell line such as MM6. Accuracy and specificity of these test methods are comparable, but in general methods using whole blood, PBMC and the MM6 cell line appear to perform best (Hoffmann et al, 2005b).

Table 9.3.1 summarises the performance of *in vitro* methods presented in the five BRDs and Table 9.3.2 compares the *in vivo* and *in vitro* pyrogen tests regarding their strengths, weaknesses, costs, time, limitations.

However, most studies (as this one) are done with model pyrogens and as yet little experience is available in the field, e.g. as part of the final batch release test-package. Experience and thus confidence in these methods will grow once regulatory authorities approve these methods and more manufacturers start to employ them. Then, on a case by case situation, it should be determined which method is best suited for the actual situation and demonstrates to pick out the appropriate, i.e. pyrogenic batches of the medicinal product.

BRD: MM6/IL-6 March, 2006

Table 9.3.1: Summary of the performance of in vitro pyrogen tests based on monocytoid cells (see Tables 7.2.2; 7.2.4; 6.1.3)

Test	System	Read-out	Intralaboratory reproducibility (%)	Interlaboratory reproducibility (%)	Sensitivity (%)	Specificity (%)
WB/IL-6	whole blood	IL-6	DL: 83.3 NL1: 94.4 NL2: 100	DL-NL1: 85.4 DL-NL2: 85.4 NL1-NL2: 92.0	88.9	96.6
WB/IL-1	whole blood	IL-1β	DL: 88. 9 NL1: 95.8 NL2: 94.4	DL-NL1: 72.9 DL-NL2: 81.6 NL1-NL2: 70.2	72.7	93.2
96-wells WB/IL-1 [1]	whole blood	IL-1β	-	DL-NL1: 88.1 DL-NL2: 89.7 NL1-NL2: 91.5	98.8	83.6
CRYO WB/Il-1	cryo whole blood	IL-1β	-	DL-NL1: 91.7 DL-NL2: 91.7 NL1-NL2: 91.7	97.4	81.4
KN CRYO WB/Il-1 [2]	cryo whole blood	IL-1β	-	DL-NL1: 83.3 DL-NL2: 100 NL1-NL2: 83.3	88.9	94.4
PBMC/IL6	PBMC	IL-6	DL: 94.4 NL1: 100 NL2: 94.4	DL-NL1: 84.0 DL-NL2: 86.0 NL1-NL2: 90.0	92.2	95.0
PBMC-CRYO/IL-6 [3]	PBMC	IL-6	-	DL-NL1: 96 DL-NL2: 76 NL1-NL2: 80	93.3	76.7
MM6/IL-6	MM6	IL-6	DL: 100 NL1: 94.4 NL2: 94.4	DL-NL1: 90.0 DL-NL2: 89.6 NL1-NL2: 83.3	95.5	89.8

DL = developing laboratory; NL1, NL2 = naive laboratory 1 and 2
1 = data provided in Section 13 of WB/IL-1 BRD
2 = data provided in Section 13 of CRYO WB/IL-1 BRD
3 = data provided in Section 13 of PBMC/IL-6 BRD

Table amended from Hoffmann et al 2005b; results with THP cells not included

BRD: MM6/IL-6 March, 2006

Table 9.3.2: Comparison of the in vivo and in vitro pyrogen tests regarding their strengths, weaknesses, costs, time, limitations

	Rabbit pyrogen test	**BET / LAL**	**In vitro pyrogen test**
Test materials	Liquids	Clear liquids	Liquids, potentially cell preparations, solid materials
Pyrogens covered	All (possible species differences to humans for non-endotoxin pyrogens)	Endotoxin from Gram-negative bacteria	(probably) all
Limit of detection (LPS)	0,5 EU	0,1 EU (some variants down to 0,01 EU)	0,5 EU (validated PM), some variants down to 0,001 EU
Ethical concerns	Animal experiment	About 10% lethality to bled horseshoe crabs	Some assays: blood donation
Costs*	High (200-600$/sample)	Low (50-150$/sample)	Medium (100-350$/sample)
Time required	27 h	45 min	24-30h**
Materials not testable	Short-lived radiochemicals, anesthetics, sedatives, analgetics, chemotherapeutics, immunomodulators, cytokines, corticosteroids	Most biologicals, glucan-containing preparations (herbal medicinal products, cellulose-filtered products), lipids, microsomes, cellular therapeutics	Not known (some of the materials not testable in rabbits require adaptations)
Others	No positive or negative control included, strain differences, stress affects body temperature	Potency of LPS from different bacterial species in mammals not reflected, false-positive for glucans	Possible donor differences, need to exclude hepatitis/HIV and acute infections / allergies of donors, dedifferentiation of cell lines

* = We consulted the laboratories participating in the validation study and a consultant regarding the costs of the tests. The figures we received vary significantly depending on the facility (e.g. industry, contract laboratory, control authority), frequency of testing, specific test requirements, country, etc.

** = interference testing might increase duration by 24 hours

10 Animal Welfare Considerations (Refinement, Reduction, and Replacement)

10.1 Diminish animal use

Describe how the proposed test method will refine (reduce or eliminate pain or distress), reduce, or replace animal use compared to the reference test method.

Depending on the medicinal product, one of two animal-based pyrogen tests is currently prescribed by the respective Pharmacopoeias, i.e. the rabbit pyrogen test and the bacterial endotoxin test (BET). The rabbit pyrogen test detects various pyrogens but alone the fact that large numbers of animals are required to identify a few batches of pyrogen-containing samples argues against its use. In the past two decades, the declared intention to refine, reduce and replace animal testing, has lowered rabbit pyrogen testing by 80% by allowing to use the BET as an alternative pyrogen test for certain medicinal products. Bacterial endotoxin is the pyrogen of major concern to the pharmaceutical industry due to its ubiquitous sources, its stability and its high pyrogenicity. With the BET, endotoxin is detected by its capacity to coagulate the amoebocyte lysate from the haemolymph of the American horseshoe crab, *Limulus polyphemus*, a principle recognised some 40 years ago (Levin and Bang, 1964). In the US, *Limulus* crabs are generally released into nature after drawing about 20% of their blood and therefore most of these animals survive. However, the procedure still causes mortality of about 30,000 horseshoe crabs per year, which adds to more efficient threats of the horseshoe crab population like its use as bait for fisheries, habitat loss and pollution.

The proposed MM6/IL-6 test method is an alternative for the rabbit test and the BET. By replacing the rabbit test or the BET, the lives of rabbits and horseshoe crabs are spared.

10.2 Continuation of animal use

If the proposed test method requires the use of animals, the following items should be addressed:

10.2.1 *Describe the rationale for the need to use animals and describe why the information provided by the proposed test method requires the use of animals (i.e., cannot be obtained using non-animal methods).*

For the culturing of the MM6 cells, foetal bovine serum is used in the culture medium. Though alternative serum-replacing cell culture reagents are currently available, there are not sufficient data to substantiate their compliance with the present method. It should be confirmed that the monocytoid cells proliferate and especially respond similar when challenged with pyrogens in the absence and presence of bovine serum.

10.2.2 *Include a description of the sources used to determine the availability of alternative test methods that might further refine, reduce, or replace animal use for this testing. This should, at a minimum, include the databases searched, the search strategy used, the search date(s), a discussion of the results of the search, and the rationale for not incorporating available alternative methods.*

Not applicable.

10.2.3 *Describe the basis for determining that the number of animals used is appropriate.*

Not applicable.

10.2.4 *If the proposed test method involves potential animal pain and distress, discuss the methods and approaches that have been incorporated to minimize and, whenever possible, eliminate the occurrence of such pain and distress.*
Not applicable.

11 Practical Considerations

11.1 Transferability

Discuss the following aspects of proposed test method transferability. Include an explanation of how this compares to the transferability of the in vivo reference test method and, if applicable, to a comparable validated test method with established performance standards.

In general, the proposed test method is not unlike other bioassays and immunoassays that are performed routinely in many laboratories.

11.1.1 *Discuss the facilities and major fixed equipment needed to conduct a study using the proposed test method*.

No extraordinary facilities are required. General laboratory equipment for aspeptic operations, e.g. cell culture (laminar airflow cabinets), and analytical instruments for performing immunoassays, e.g. microtiter plate reader and –washer, are sufficient to perform the proposed test method.

11.1.2 *Discuss the general availability of other necessary equipment and supplies*.

All supplies and reagents are readily available on the market. In contrast, availability of sufficient rabbits of adequate weight and in good health for the *in vivo* reference test is sometimes reported a limitation.

The MM6/IL-6 assay is not patented. The cell line used, the MM6, was established by Prof. H.W.L. Ziegler-Heitbrock, Institute for Immunology, University of Munich, Munich, Germany. It can be obtained for research purposes only from Prof. H.W.L. Ziegler-Heitbrock or from the German Collection of Microorganisms and Cell Cultures (DSMZ), Braunschweig, Germany.

The conditions for licensing of the cell line are to be negotiated individually with Prof. Ziegler-Heitbrock.

11.2 Training

Discuss the following aspects of proposed test method training. Include an explanation of how this compares to the level of training required to conduct the in vivo reference test method and, if applicable, a comparable validated test method with established performance standards.

11.2.1 *Discuss the required level of training and expertise needed for personnel to conduct the proposed test method.*

The proposed test method requires personnel trained for general laboratory activities in cell biology and immunochemistry or biochemistry. Techniques they should master are not unlike cell culture (aseptic operations) and immunological techniques (especially ELISA). Such expertise is available in most if not all QC-laboratories.

11.2.2 *Indicate any training requirements needed for personnel to demonstrate proficiency and describe any laboratory proficiency criteria that should be met.*

Personnel should demonstrate that they master the execution of the test. The candidate should demonstrate to meet all the appropriate assay acceptance criteria and yield accurate results (outcome) using selected test items.

11.3 Cost Considerations

Discuss the cost involved in conducting a study with the proposed test method. Discuss how this compares to the cost of the in vivo reference test method and, if applicable, with that of a comparable validated test method with established performance standards.

Three factors contribute to the cost of the proposed test method: availability of monocytoid cells, cost of the reagents for the immunoassay and, last but not least, personnel.

The conditions (costs) for licensing of the MM6 cell-line are to be negotiated individually with Prof. Ziegler-Heitbrock (see section 11.1.2).

Since the proposed test method is relatively more labor-intensive, it is estimated that the cost of the proposed test method is higher than those for the BET or the *in vivo* reference test using rabbits. Obviously, a higher throughput of tests (runs/year) such as in a QC-laboratory of a multi-product facility or in a contract research organisation will significantly reduce the costs per assay.

However, especially with pharmaceuticals of biological origin, the proposed test method may be most cost-effective, since these products all too often are incompatible with the BET and by their nature preclude the re-use of the rabbits.

11.4 Time Considerations

Indicate the amount of time needed to conduct a study using the proposed test method and discuss how this compares with the in vivo reference test method and, if applicable, with that of a comparable validated test method with established performance standards.

Essentially the test stretches two working days. On day one the testing materials are prepared and incubated overnight with the monocytoid cells. On the second day the amount of excreted cytokines is determined by immunoassay. Total time from start to result are 24 hours.

It is thus concluded that the proposed test method will take more time than the reference tests, either the rabbit test or the BET. It should be noted that rabbits are tested prior to their first use by a sham test.

12 References

List all publications referenced in the submission.

References in bold are included as hardcopies in Appendix B

Barnett, V., Lewis, T. (1984). Outliers in statistical data. 2nd edition, Chichester, US: John Wiley & Sons, pp 171-172.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Clopper, C. J., Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26: 404-413

Co Tui & Schrift, M.H. (1942). A tentative test for pyrogen in infusion fluids. Proc. Soc. Exp. Biol. Med. 49: 320-330.

Cooper JF, Levin J, Wagner Jr. HN (1971). Quantitative comparison of in vitro and *in vivo* methods for the detection of endotoxin. J Lab Clin Med 78: 138-145.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005a). Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Pertbauer A, Schindler S, et al. (2005b) Validation of novel pyrogen tests based on human monocytoid cell lines. J Immunol Methods, 298(1-2):161-173.

Hothorn, L. A. (1995). Biostatistical analysis of the 'control vs. k treatments' design including a positive control group. In: Testing principls in clinical and preclinical trials, pp 19-26, Stuttgart, Germany: Gustav Fischer Verlag.

Levin J, Bang FB (1964). A description of cellular coagulation in the Limulus. Bull Johns Hopkins Hosp 155: 337-345.

Nakagawa, Y., Maeda, H. and Murai, T. (2002). Evaluation of the *in vitro* pyrogen test system based on proinflammatory cytokine release from human monocytes: comparison with a human whole blood culture test system and with the rabbit pyrogen test. Clin. Diagn. Lab. Immunol. 9: 588-592

Pasare, C. and Medzhitov, R. (2003). Toll pathway-dependent blockade of CD4+CD25+ T cell-mediated suppression by dendritic cells. Science 299: 1033-1036.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, November 1989: 17-18.

Poole, S. and Gaines Das, R.E. (2001). Towards a 'human pyrogen test'. J. Parenter. Sci. 6: 63-66.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' *in vitro* test for pyrogens. J Immunol Methods 274: 209-220.

Reddi, K., Nair, S.P. et al (1996). Surface-associated material from the bacterium Actinobacillus actinomycetemcomitans contains a peptide which, in contrast to lipopolysaccharide, directly stimulates fibroblast interleukin-6 gene transcription. Eur. J. Biochem. 236: 871-876.

Snedecor, G. W., Cochran, W. G. (1989). Statistical methods, 8th edition, Ames, US: The Iowa State University Press.

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

13 Supporting Materials (Appendices)

13.1 Standard operating procedure (SOP) of the proposed method

Provide the complete, detailed protocol for the proposed test method.

Appendix A includes the original protocol of the MM6/IL-6 test (*Detailed protocol MM6/IL-6: "In vitro pyrogen test using MONOMAC 6 CELLS"; electronic file name: SOP MM6-IL-6*), the protocol used in the validation study (*marked with internal identifier SopMM6v08; electronic file name SOP MM6-IL-6 validation)* and the trial plan of the validation study.

13.2 Standard operating Procedure (SOP) of the reference method

Provide the detailed protocol(s) used to generate reference data for this submission and any protocols used to generate validation data that differ from the proposed protocol.
Not applicable.

13.3 Publications

Provide copies of all relevant publications, including those containing data from the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method with established performance standards.

Not all of the publications cited in the BRD (see section 12) are included as hardcopies in Appendix B, e.g. publications on statistical methods are not given.

Remark: The same set of hardcopies was included into Appendix B of all of the 5 submitted BRDs. Therefore some of the publications in Appendix B might not be referenced in the current BRD nor included in the list of the publications in section 12. Three general publications were added, which are not cited in any BRD but might be useful as background information to the validation study: the ECVAM publications on validation (Balls et al, 1995; Curren et al, 1995; Hartung et al, 2004). Several publications were included, which either give more background information on the human fever reaction or report on specific studies using *in vitro* pyrogen tests.

List of hard copies

Andrade SS, Silveira RL, Schmidt CA, Junior LB, Dalmora SL. (2003) Comparative evaluation of the human whole blood and human peripheral blood monocyte tests for pyrogens. Int J Pharm. Oct 20;265(1-2):115-24.

Balls et al. (1995) Practical aspects of the validation of toxicity test procedures. ECVAM Workshop Report 5. ATLA 23, 129-147.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Carlin G, Viitanen E. (2003) In vitro pyrogenicity of a multivalent vaccine: Infanrix. PHARMEUROPA, 15(3), 418-423.

Curren et al (1995) The Role of prevalidation in the development, validation and acceptance of Alternative Methods. ECVAM Prevalidation Task Force Report 1. ATLA 23, 211-217.

De Groote, D., Zangerle, P.F., Gevaert, Y., Fassotte, M.F., Beguin, Y., Noizat-Pirenne, F., Pirenne, J., Gathy, R., Lopez, M., Dehart I., (1992). Direct stimulation of cytokines (IL-1 beta, TNF-alpha, IL-6, IL-2, IFN-gamma and GM-CSF) in whole blood. I. Comparison with isolated PBMC stimulation. Cytokine 4, 239.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Fennrich S, Wendel A and Hartung T. (1999). New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 16:146-149

Fennrich S, Fischer M, Hartung T, Lexa P, Montag-Lessing T, Sonntag H-G, Weigandt M und Wendel A. (1999). Detection of endotoxins and other pyrogens using human whole blood. Dev. Biol. Standards 101:131-139

Gaines Das et al (2004). Monocyte activation test for pro-inflammatory and pyrogenic contaminants of parenteral drugs: test design and data analysis. J. of Immunological Methods 288, 165-177.

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hartung et al (2004) A modular approach to the ECVAM principles on test validity. ATLA 32, 467-472.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005). Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Peterbauer A, Schindler S, et al (2005). International validation of novel pyrogen tests based on human monocytoid cells. J Immunol Methods 298: 161-173.

Jahnke M, Weigand, Sonntag H-G. (2000). Comparative testing for pyrogens in parenteral drugs using the human whole blood pyrogen test, the rabbit in vivo pyrogen test and the LAL test. European Journal of Parenteral Science 5(2):39-44

Nakagawa Y, Maeda H, Murai T. (2002) Evaluation of the in vitro pyrogen test system based on proinflammatory cytokine release from human monocytes: comparison with a human whole blood culture test system and with the rabbit pyrogen test. Clin Diagn Lab Immunol. May;9(3):588-97.

Pasare, C. and Medzhitov, R. (2003). Toll pathway-dependent blockade of CD4+CD25+ T cell-mediated suppression by dendritic cells. Science 299: 1033-1036.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1998) The detection of pyrogens in blood products using an ex vivo whole blood culture assay. J Immunoassay. May-Aug;19(2-3):95-111.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1999) Differentiation between endotoxin and non-endotoxin pyrogens in human albumin solutions using an ex vivo whole blood culture assay. J Immunoassay. Feb-May; 20(1-2):79-89.

Poole S, Thorpe R, Meager A, Hubbard AJ, Gearing AJ (1988). Detection of pyrogen by cytokine release. Lancet 8577, 130.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, November 1989: 17-18.

Poole S, Musset MV (1989). The International Standard for Endotoxin: evaluation in an international collaborative study. J Biol Stand 17: 161-171.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' in vitro test for pyrogens. J Immunol Methods 274: 209-220.

Ray A, Redhead K, Selkirk S, Polle S (1991) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211-218.

Reddi, K., Nair, S.P. et al (1996). Surface-associated material from the bacterium Actinobacillus actinomycetemcomitans contains a peptide which, in contrast to lipopolysaccharide, directly stimulates fibroblast interleukin-6 gene transcription. Eur. J. Biochem. 236: 871-876.

Schindler S, Bristow A, Cartmell T, Hartung T and Fennrich S. (2003). Comparison of the reactivity of human and rabbit blood towards pyrogenic stimuli. ALTEX 20:59-63.

Schindler et al (2004) Cryopreservation of human whole blood for pyrogenicity testing. Journal of Immunological Methods 294, 89–100

Schins RP, van Hartingsveldt B, Borm PJ. Ex vivo cytokine release from whole blood. A routine method for health effect screening. Exp Toxicol Pathol. 1996 Nov; 48(6):494-6.

Spreitzer I, Fischer M, hartzsch K, Lüderitz-Püschel U and Montag T. (2000). Comparative Study of Rabbit Pyrogen Test and Human whole Blood Assay on Human Serum Albumin. ALTEX 19 (Suppl. 1):73-75

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

13.4 Original data

Include all available non-transformed original data for both the proposed test method, the in vivo reference test method, and if applicable, a comparable validated test method with established performance standards.

NOTE: The original data of the ELISA-plate reader were collected by S.Hoffman and ECVAM. These are available on the CD which goes with the BRD.

13.5 Performance standards

If appropriate performance standards for the proposed test method do not exist, performance standards for consideration by NICEATM and ICCVAM may be proposed. Examples of established performance standards can be located on the ICCVAM / NICEATM web site at http://iccvam.niehs.nih.gov.

BRD: MM6/IL-6 March, 2006

APPENDIX A

Trial plan "Comparison And Validation Of Novel Pyrogen Tests Based On The Human Fever Reaction" Acronym: Human (e) Pyrogen Test

Detailed protocol MM6/IL-6: "In vitro pyrogen test using MONOMAC 6 CELLS" *marked with internal identifier SopMM6v08* – used in validation study (electronic file name: SOP MM6-IL-6 validation)

Original protocol MM6/IL-6: "In vitro pyrogen test using MONOMAC 6 CELLS" (electronic file name: SOP MM6-IL-6)

BRD: MM6/IL-6 March, 2006

APPENDIX B

Not all of the publications cited in the BRD (see section 12) are included as hardcopies in Appendix B, e.g. publications on statistical methods are not given.

Remark: The same set of hardcopies was included into Appendix B of all of the 5 submitted BRDs. Therefore some of the publications in Appendix B might not be referenced in the current BRD nor included in the list of the publications in section 12. Three general publications were added, which are not cited in any BRD but might be useful as background information to the validation study: the ECVAM publications on validation (Balls et al, 1995; Curren et al, 1995; Hartung et al, 2004). Several publications were included, which either give more background information on the human fever reaction or report on specific studies using *in vitro* pyrogen tests.

List of hard copies

Andrade SS, Silveira RL, Schmidt CA, Junior LB, Dalmora SL. (2003) Comparative evaluation of the human whole blood and human peripheral blood monocyte tests for pyrogens. Int J Pharm. Oct 20;265(1-2):115-24.

Balls et al. (1995) Practical aspects of the validation of toxicity test procedures. ECVAM Workshop Report 5. ATLA 23, 129-147.

Beutler B, Rietschel ET (2003). Innate immune sensing and its roots: the story of endotoxin. Nature Rev Immunol. 3: 169-176.

Bleeker, W.K., de Groot, E.M., den Boer, P.J. et al (1994). Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol 22: 835-840.

Carlin G, Viitanen E. (2003) In vitro pyrogenicity of a multivalent vaccine: Infanrix. PHARMEUROPA, 15(3), 418-423.

Curren et al (1995) The Role of prevalidation in the development, validation and acceptance of Alternative Methods. ECVAM Prevalidation Task Force Report 1. ATLA 23, 211-217.

De Groote, D., Zangerle, P.F., Gevaert, Y., Fassotte, M.F., Beguin, Y., Noizat-Pirenne, F., Pirenne, J., Gathy, R., Lopez, M., Dehart I., (1992). Direct stimulation of cytokines (IL-1 beta, TNF-alpha, IL-6, IL-2, IFN-gamma and GM-CSF) in whole blood. I. Comparison with isolated PBMC stimulation. Cytokine 4, 239.

Dinarello CA (1999). Cytokines as endogenous pyrogens. J Infect Dis 179 (Suppl 2): S294-304.

Dinarello CA, O'Connor JV, LoPreste G, Swift RL (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J Clin Microbiol 20: 323-329.

Duff GW, Atkins E (1982). The detection of endotoxin by in vitro production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52: 323-331.

Fennrich S, Wendel A and Hartung T. (1999). New applications of the human whole blood pyrogen assay (PyroCheck). ALTEX 16:146-149

Fennrich S, Fischer M, Hartung T, Lexa P, Montag-Lessing T, Sonntag H-G, Weigandt M und Wendel A. (1999). Detection of endotoxins and other pyrogens using human whole blood. Dev. Biol. Standards 101:131-139

Gaines Das et al (2004). Monocyte activation test for pro-inflammatory and pyrogenic contaminants of parenteral drugs: test design and data analysis. J. of Immunological Methods 288, 165-177.

Greisman SE, Hornick RB (1969). Comparative pyrogenic reactivity of rabbit and man to bacterial endotoxin. Proc Soc Exp Biol Med 131: 1154-1158.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15: 425–433.

Hartung T, Wendel A (1996). Detection of pyrogens using human whole blood. In Vitro Toxicol 9: 353-359.

Hartung T, Aaberge I, Berthold S et al (2001). Novel pyrogen tests based on the human fever reaction. Altern Lab Anim 29: 99-123.

Hartung et al (2004) A modular approach to the ECVAM principles on test validity. ATLA 32, 467-472.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T. (2005). Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modeling. Journal of Endotoxin Research 11(1): 26-31

Hoffmann S, Peterbauer A, Schindler S, et al (2005). International validation of novel pyrogen tests based on human monocytoid cells. J Immunol Methods 298: 161-173.

Jahnke M, Weigand, Sonntag H-G. (2000). Comparative testing for pyrogens in parenteral drugs using the human whole blood pyrogen test, the rabbit in vivo pyrogen test and the LAL test. European Journal of Parenteral Science 5(2):39-44

Nakagawa Y, Maeda H, Murai T. (2002) Evaluation of the in vitro pyrogen test system based on proinflammatory cytokine release from human monocytes: comparison with a human whole blood culture test system and with the rabbit pyrogen test. Clin Diagn Lab Immunol. May;9(3):588-97.

Pasare, C. and Medzhitov, R. (2003). Toll pathway-dependent blockade of CD4+CD25+ T cell-mediated suppression by dendritic cells. Science 299: 1033-1036.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1998) The detection of pyrogens in blood products using an ex vivo whole blood culture assay. J Immunoassay. May-Aug;19(2-3):95-111.

Pool EJ, Johaar G, James S, Petersen I, Bouic P. (1999) Differentiation between endotoxin and non-endotoxin pyrogens in human albumin solutions using an ex vivo whole blood culture assay. J Immunoassay. Feb-May; 20(1-2):79-89.

Poole S, Thorpe R, Meager A, Hubbard AJ, Gearing AJ (1988). Detection of pyrogen by cytokine release. Lancet 8577, 130.

Poole S, Thorpe R, Meager A, Gearing AJ (1988). Assay of pyrogenic contamination in pharmaceuticals by cytokine release from monocytes. Dev Biol Stand 69: 121-123.

Poole S, Selkirk S, Rafferty B, Meager A, Thorpe R, Gearing A (1989). Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of

the European Workshop on detection and quantification of pyrogen. Pharmeuropa special Vol 1, November 1989: 17-18.

Poole S, Musset MV (1989). The International Standard for Endotoxin: evaluation in an international collaborative study. J Biol Stand 17: 161-171.

Poole S, Mistry Y, Ball C et al (2003). A rapid 'one-plate' in vitro test for pyrogens. J Immunol Methods 274: 209-220.

Ray A, Redhead K, Selkirk S, Polle S (1991) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211-218.

Reddi, K., Nair, S.P. et al (1996). Surface-associated material from the bacterium Actinobacillus actinomycetemcomitans contains a peptide which, in contrast to lipopolysaccharide, directly stimulates fibroblast interleukin-6 gene transcription. Eur. J. Biochem. 236: 871-876.

Schindler S, Bristow A, Cartmell T, Hartung T and Fennrich S. (2003). Comparison of the reactivity of human and rabbit blood towards pyrogenic stimuli. ALTEX 20:59-63.

Schindler et al (2004) Cryopreservation of human whole blood for pyrogenicity testing. Journal of Immunological Methods 294, 89–100

Schins RP, van Hartingsveldt B, Borm PJ. Ex vivo cytokine release from whole blood. A routine method for health effect screening. Exp Toxicol Pathol. 1996 Nov; 48(6):494-6.

Spreitzer I, Fischer M, hartzsch K, Lüderitz-Püschel U and Montag T. (2000). Comparative Study of Rabbit Pyrogen Test and Human whole Blood Assay on Human Serum Albumin. ALTEX 19 (Suppl. 1):73-75

Taktak YS, Selkirk S, Bristow AF et al (1991). Assay of pyrogens by interleukin-6 release from monocytic cell lines. J Pharm Pharmacol 43: 578-582.

Tsuchiya S, Yamabe M, Yamaguchi Y et al (1980). Establishment and characterisation of a human acute monocytic leukemia cell line (THP-1). Int J Cancer 26 : 171-176.

Ziegler-Heitbrock HWL, Thiel E, Futterer A et al (1988). Establishment of a human cell line (MONO MAC 6) with characteristics of mature monocytes. Int J Cancer 41: 456-461.

BRD: MM6/IL-6 March, 2006

APPENDIX C

List of abbreviations and definitions

Accuracy	The ability of a test system to provide a test result close to the accepted reference value for a defined property.
BET	The bacterial endotoxin test is used to detect or quantify endotoxins of gram-negative bacterial origin using amoebycte lysate from horseshoe crab (*Limulus polyphemus* or *Tachypleus tridentatus*
BRD	Background Review Document
CRYO WB/IL-1	Whole blood assay (using cryopreserved blood) with IL-1 as endpoint
CV	coefficient of variation
DL	Developing laboratory = laboratory which developed the method or the most experienced laboratory
ELC	Endotoxin limit concentration; maximum quantity of endotoxin allowed in given parenterals according to European Pharmacopoeia
Endotoxins	Endotoxins are a group of chemically similar cell-wall structures of Gram-negative bacteria, i.e. lipopolysaccharides
ELISA	Enzyme linked immunosorbent assay
EU/ml	European Units per ml
IL-1	interleukin 1
IL-6	interleukin 6
Intralaboratory reproducibility	A determination of the extent that qualified people within the same laboratory can independently and successfully replicate results using a specific protocol at different times.
Interlaboratory reproducibility	A measure of the extent to which different qualified laboratories, using the same protocol and testing the same substances, can produce qualitatively and quantitatively similar results. Inter-laboratory reproducibility is also referred to as between-laboratory reproducibility.
KN	University of Konstanz (Konstanz, Germany), developing laboratory WB/IL-1 and CRYO WB/IL-1
LPS	lipopolysaccharides
MM6	MONO MAC-6 cell line
MM6/IL-6	In vitro pyrogen test using MM6 cell line and IL-6 release

	as an endpoint
MVD	Maximum valid dilution; the MVD is the quotient of the ELC and the detection limit
NIBSC	National Institute for Biological Standards and Control (London, UK), developing laboratory for WB/IL-6
NL	naïve laboratory = laboratory with non or minor experience with the method
NPC	negative product control (clean, released lot of the nominated product under test)
Novartis	Novartis (Basel, Switzerland), developing laboratory PBMC/IL-6
OD	optical density
PBMC	Peripheral blood mononuclear cells
PBMC/IL-6	In vitro pyrogen test using fresh peripheral blood mononuclear cells and IL-6 release as endpoint
PBMC-CRYO/IL-6	In vitro pyrogen test using cryopreserved peripheral blood mononuclear cells and IL-6 release as endpoint
PEI	Paul-Ehrlich Institut (Langen, Germany), participating laboratory
PM	prediction model = is an explicit decision-making rule for converting the results of the in vitro method into a prediction of in vivo hazard
PPC	positive product control (product under test spiked with 0.5 EU/ml of WHO-LPS (code 94/580)
Prevalidation study	A prevalidation study is a small-scale inter-laboratory study, carried out to ensure that the protocol of a test method is sufficiently optimised and standardised for inclusion in a formal validation study. According to the ECVAM principles, the prevalidation study is divided into three phases: protocol refinement, protocol transfer and protocol performance (Curren et al, ATLA 23, 211-217).
Pyrogens	fever-causing materials
Pyrogens, endogenous	endogenous pyrogens are messenger substances released by blood cells reacting to pyrogenic materials; e.g. IL-1, IL-6, TNF-α, prostaglandin E_2
Pyrogens, exogenous	exogenous pyrogens derive from bacteria, viruses, fungi or from the host himself
Reliability	Measures of the extent to which a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and interlaboratory reproducibility and intra-laboratory repeatability.

Relevance	Relevance of a test method describes whether it is meaningful and useful for a particular purpose. It is the extent to which the measurement result and uncertainty can accurately be interpreted as reflecting or predicting the biological effect of interest.
Repeatibility	Repeatability describes the closeness of agreement between test results obtained within a single laboratory when the procedure is performed independently under repeatability conditions, i.e. in a set of conditions including the same measurement procedure, same operator, same measuring system, same operating conditions and same location, and replicated measurements over a short period of time.
RIVM	National Institute of Public Health and the Environment (Bilthoven, The Netherlands), developing laboratory MM6/IL-6 method
Sensitivity	Sensitivity is the proportion of all positive/active substances that are correctly classified by a test method.
Specificity	Specificity is proportion of all negative/inactive substances that are correctly classified by a test method.
TMB	chromogenic substrate 3,3´,5,5´ -tetramethylbenzidine
TNF-α	tumour necrosis factor-α
USP	US Pharmacopoeia
Validation	Validation is the process by which the reliability and relevance of a procedure are established for a specific purpose
Validation study	A validation study is a large-scale interlaboratory study, designed to assess the reliability and relevance of an optimised method for a particular purpose
WB/IL-1	Whole blood assay (using fresh blood) with IL-1 release as endpoint
WB/IL-6	Whole blood assay (using fresh blood) with IL-6 release as endpoint
WHO	World Health Organization

[This Page Intentionally Left Blank]

Validation of Biomedical Testing Methods

***In vitro* pyrogen test using MONOMAC 6 CELLS**

Standard Operating Procedure

Only the responsible of the GLP/QA Unit is allowed to make copies of this document. Extra examples can be obtained at the GLP/QA Unit. Quality Documents are only valid if they are signed by the responsible of the GLP/QA Unit in and provided with a blue copy number	Copy number

Identity

JOINT RESEARCH CENTRE
EUROPEAN COMMISSION
Institute for Health and Consumer Protection
ECVAM

Standard Operating procedure

In vitro pyrogen test using MONOMAC 6 TEST

Version number: 4
Applicable from: 23 04 02
Expired at:

Drafted by:	Name	Stephen Poole
	Date	26-09-01
	Signature	
Reviewed by:	Name	Ria Nibbeling
	Date	28-02-02
	Signature	
Reviewed by:	Name	Ria Nibbeling
	Date	23-04-02
	Signature	
	Name	
	Date	
	Signature	

*Owner/Trainer: Signature: Date:

PAGE OF CHANGES

Date of change/ Date of draft:	Version– number:	Changed page(s):	Summary of the change(s):	Changed by/Sign.:
280202	2		Overall revision	RN
140402	3	12	Pre-incubation of cells	RN
230402	4	15	Chapter 8	RN

TABLE OF CONTENTS

Page No.

1 ***INTRODUCTION*** ***5***
2 ***PURPOSE*** ***7***
3 ***SCOPE / LIMITATIONS*** ***7***
4 ***METHOD OUTLINE*** ***7***
5 ***DEFINITIONS / ABBREVIATIONS*** ***7***
6 ***MATERIALS*** ***9***
7 ***METHODS*** ***11***
8 ***DATA ANALYSIS, PREDICTION MODEL AND RELATED ERRORS*** ***17***
9 ***HEALTH SAFETY AND ENVIRONMENT*** ***18***
10 REFERENCE
11ANNEX ***19***

1. INTRODUCTION

Parenteral pharmaceutical products must be shown to be free from pyrogenic (fever–inducing) contamination. While a pyrogen may in general be defined as any substance that causes fever, the pyrogens that almost invariably contaminate parenteral pharmaceuticals are bacterial endotoxins (lipopolysaccharides, LPS) from Gram–negative bacteria (Mascoli and Weary, 1979a, 1979b). There are two Pharmacopoeial tests for pyrogenic contamination: the rabbit pyrogen test and the Limulus amoebocyte lysate (LAL) test. The rabbit pyrogen test, which detects LPS and other pyrogens, involves measuring the rise in body temperature evoked in rabbits by the intravenous injection of a sterile solution of the substance to be examined. In contrast, the LAL test detects only LPS: it is described in Pharmacopoeias as the bacterial endotoxins test (BET). The principle of the LAL–test is that LPS causes extracellular coagulation of the blood (haemolymph) of the horseshoe crab, *Limulus polyphemus.* (Levin & Bang, 1964). Although the LAL test is gradually superseding the rabbit pyrogen test, hundreds of thousands of rabbit pyrogen tests are still carried out each year around the world, largely on products which cannot, for one reason or another, be tested in the LAL test. While proving generally reliable, both the rabbit pyrogen test and LAL test have shortcomings. The rabbit pyrogen test uses experimental animals, is costly and is not quantitative. The LAL test gives false negatives with certain products, can overestimate the pyrogen content of other products and does not detect pyrogens other than bacterial endotoxin (LPS), such as Gram–positive exotoxins, viruses and fungi (Dinarello et al., 1984; Poole et al., 1988; Ray et al., 1990; Taktak et al., 1991; Fennrich et al., 1999).

The basis of the rabbit pyrogen test is the *in vivo* stimulation by exogenous pyrogens (usually LPS) of rabbit peripheral blood monocytes to produce the endogenous pyrogens that cause fever. The endogenous pyrogens are pyrogenic cytokines such as tumour necrosis factorα (TNFα), interleukin–1 (IL–1α and IL–1β, two separate gene products), IL–6 and IL–8 (Dinarello et al., 1999). In view of the shortcomings of the rabbit pyrogen test and the LAL test, in vitro pyrogen tests that utilise the exquisite sensitivity to exogenous pyrogen of monocytes have been proposed. In such tests, products are incubated with human peripheral blood monocytes (or mononuclear cells, PBMNC CELLS, or leukocytes) and the conditioned media assayed for pyrogenic cytokines (Duff & Atkins, 1982; Dinarello et al., 1984; Poole et al., 1988; Poole, 1989; Hansen and Christensen, 1990; Taktak et al., 1991; Bleeker et al., 1994).

The isolation of monocytes/leukocytes from whole blood is labour–intensive and time–consuming, technically sophisticated, requires expensive reagents and does not guarantee the isolation of cells in a non–activated state. This prompted the evaluation of various cell lines which retain monocytic characteristics, including the capacity to synthesise and secrete pyrogenic cytokines (Taktak et al., 1991). Taktak *et al*. measured LPS–induced IL–1β and IL–6 release from MONO MAC 6 cells and THP–1 cells and concluded that IL–6 release by MONO MAC 6 cells was the most appropriate readout for an *in vitro* pyrogen test ('monocyte test') because immunoreactive IL–6, unlike immunoreactive IL–1 and TNFα, is secreted entirely into the cell–conditioned medium in large quantities, permitting its complete estimation. The test was applied to three batches of therapeutic human serum albumin (HSA) that had caused adverse reactions in recipients. The MONO MAC 6/IL–6 test

Institute for Health and Consumer Protection

detected pyrogenic contamination in the HSA that had not been detected in the rabbit pyrogen test and the LAL test.

2. PURPOSE

To develop an *in vitro* pyrogen test that will serve as a replacement for the rabbit pyrogen test.

3. SCOPE / LIMITATIONS

The method described below is for the evaluation of an in vitro MONOMAC 6/IL–6 release test. It is not a 'finalised' test system for the testing of medicinal products. The method may be applied only to preparations that have been validated with the method, i.e. shown not to interfere in the test system: see Heading 8.

4. METHOD OUTLINE

MONOMAC 6 cells are stimulated for 16–24h with standard endotoxin (LPS) and preparations under test. Following this stimulation, the concentration of IL–6 in the cell–conditioned medium is quantified using a specific ELISA (which is calibrated in terms of the appropriate international standard). The construction of a dose–response curve for endotoxin standard versus OD-value of IL–6 ELISA permits the estimation of the pyrogenic contamination of the preparations under test. The contamination is measured in endotoxin–equivalent units.

5. DEFINITIONS / ABBREVIATIONS

µg	microgram
µl	microlitre
Ab	antibody
BSA	bovine serum albumin
CO_2	Carbon dioxide
°C	degrees Celsius (Centigrade)
DMSO	Dimethylsulfoxide
D–R	dose–response
E. coli	Escherichia coli
ELISA	Enzyme–linked immunosorbent assay
EP	European Pharmacopoeia
EU	endotoxin units
FDA	Food and Drug Administration (USA)
g	gram
h	hour
HIFCS	heat–inactivated (+56°C for 30 min) foetal calf serum
H_2O_2	hydrogen peroxide
H_2SO_4	sulphuric acid
HSA	Human Serum Albumin
IL	interleukin
IS	international standard
IU	international unit
l	litre
KOH	potassium hydroxide
LAL	limulus amoebocyte lysate
LPS	lipopolysaccharide
M	molar
MAb	monoclonal antibody
mg	milligram
min	minute
ml	millilitre
mM	millimolar
NaCl	sodium chloride
NaOH	sodium hydroxide
$NaHCO_3$	sodium hydrogen carbonate
NaH_2PO_4	sodium di–hydrogen phosphate
Na_2HPO_4	di–sodium hydrogen orthophosphate
No.	number
nm	nanometre
OD	optical density
PBS	Dulbecco's phosphate buffered saline
PC	Personal Computer
PF	pyrogen–free (items purchased as sterile and pyrogen–free or baked at 250°C for 30–60 min.)
POD	horseradish peroxidase conjugate
R	endotoxin standard
rpm	rounds per minute
RPMI	RPMI 1640 cell culture medium
RSE	Reference Standard Endotoxin

RT	room temperature
TMB	Tetramethyl benzidine
S	test sample
STD	standard
UNK	unknown
USP	United States Pharmacopoeia
x g	x gravity

6. MATERIALS

6.1. Cell line

The cell line used is human monocytic cell line MonoMac-6 (obtained from Prof. H.W.L. Ziegler-Heitbrock (Institute for Immunology, University of Munich, Munich, Germany) A Master Cell Bank and a Working Cell Bank was established at the NIBSC. Cells can be obtained from this laboratory.

6.2. Technical equipment

Incubator (37°C ± 2 °C, 5% ± 0.5 % CO_2, humidified)
Inverted microscope
Haematocytometer
Laminar flow clean bench (Class II)
Laboratory centrifuge (refrigerated) suitable for 50 ml tubes
Laboratory centrifuge suitable for 1 ml vials
Water bath (adjustable to 37 °C, 56°C and 70°C)
pH meter
ELISA-reader
Platewasher
Pipettes adjustable to 2-20 µl, 50-200 µl, 20-100 µl or 200-1000 µl (e.g. Gilson)
12-channel or 8-channel pipette
Vortex mixer

6.3. Other materials

All materials must be sterile and pyrogen free.

Tissue culture flasks, 25 cm^2, 75 cm^2 and 150 cm^2 (e.g. Costar)
Centrifuge tubes, 15 ml and 50 ml (e.g.Greiner or Falcon)
pyrogen free tips (e.g. Greiner)
96-wells tissue culture plates with lid (e.g.Costar)
polystyrene tubes (e.g.Greiner or Falcon)
plate sealers, non toxic (e.g.Dynatech Laboratories)
cryotubes, 2 ml (e.g. Nunc)
serological pipettes (5ml,10ml, 25ml, e.g. Beckton Dickinson Labware)
0.22 µm sterile filters (MilliPak 60, Millipore)

6.4 Chemicals and culture media

Sterile, pyrogen–free phosphate buffered saline (e.g. Life Technologies)
Hydrochloric acid, 0.1M, sterile filtered (Sigma, H–9892)
Pyrogen free sodiumchloride 0.9% (i.e. saline)(e.g. NPBI)
Water for injections (e.g. NPBI)
Trypan blue stain (e.g Sigma)

RPMI 1640 (e.g. Lifetechnologies, Gibco BRL, code 31870–025)
Hepes (e.g. 1M, Gibco)
Insulin (Sigma, code 1–4011)
L–Glutamine (e.g. 200 mM , Gibco)
Oxaloacetic acid (e.g. Sigma)
Sodium pyruvate (e.g. 100 mM, Gibco)
MEM non–essential amino acid solution (e.g. Gibco)
Fetal Bovine Serum (FBS) (e.g. Myoclone Super Plus FBS, Gibco, code 16000-036).
Note: The endotoxin content must be < 0.1 ng/ml, checked by LAL, after heating for 30 min. at 70°C.
Dimethyl sulphoxide (e.g. Merck)
Penicillin/Streptomycin, (10,000 IU/ml penicillin, 10 mg/ml streptomycin (e.g. Gibco, code 1514-0114))
International Standard for Endotoxin, 10,000 IU per vial (NIBSC, code 94/580)
International Standard for IL-6, 1µg/ 100000 IU per ampoule (NIBSC, code 89/548)

(HSA, 1% (a dilution in sterile PF saline of clinical grade HSA, 4.5%))
(Human serum AB (e.g. Sigma))

All other consumables are purchased as sterile and pyrogen–free and other reagents are pro analysis grade.

Institute for Health and Consumer Protection

7. METHODS

7.1. STEPS PRIOR TO CELL–CULTURE

Steps marked (^{at}) are carried out in a Class 2 laminar flow sterile cabinet, using aseptic technique and reagents and consumables that are sterile and pyrogen–free.

Preparation of aliquots of the LPS (endotoxin) standard (STD)^{at}

Make aliquots of the LPS STD: take a vial of the current IS for endotoxin (vial code 94/580, 10000 IU = EU/vial, infinite shelf life when stored at –20°C or below), and reconstitute the contents of the vial with 5 ml pyrogen free water and vortex for 30 min. This gives a 2,000 IU/ml stock solution of LPS. The stock solution may be kept at 2–8°C for up to 14 days. Alternatively, aliquot LPS standard into labelled cryotubes, freeze them upright and store them at –20°C or below (shelf life = 12 months) The following primary standards (Reference Standard Endotoxins) are identical to the IS and may be substituted for it: EC6 (USP), Lot G (FDA), BRP3 (EP). Alternatively, a working standard (control standard endotoxin) calibrated against one of these primary standards may be substituted.

Preparation of aliquots of the IL–6 standard^{at}

To make aliquots of the IL–6 standard, take a vial of the IS for IL–6 (ampoule code 89/548, 1 μg/100000 IU/ampoule) and reconstitute with 1 ml of PBS + 1% BSA (or HSA – not critical). This gives a 1 μg/ml stock solution of IL–6†.
(†: concentrations of IL–6 are expressed in pg/ml rather than IU/ml to avoid confusion with IU/ml of LPS)

Aliquot the stock solution into labelled cryo–tubes (of 2 ml capacity), freeze them upright and store them in this frozen state at –20°C or below (shelf life = six months).

A working standard, previously calibrated against the IS may be substituted for the IS. Each new batch of working standard is to be calibrated against the IS.

Preparation of medium^{at}

Prepare and store the RPMI as described by the supplier. Adjust the pH and sterile filter (0.22 μm) the medium if required.
Test a sample of the medium (from one bottle) for LPS contamination in an LAL test according to the current SOP's for LAL testing. Use the batch of medium only if the level of contamination is not greater than 0.06 IU/ml.

'Completion' of maintenance culture medium (RPMI–M)[at]

RPMI 1640 medium	500 ml
Heat–inactivated (+56°C for 30 min) foetal calf serum (HIFCS)	50 ml
L–Glutamine, 2 mM	
MEM non–essential amino acid, 0.1 mM	
Bovine insulin, 0.23 IU/ml	
Oxaloacetic acid, 1 mM	
Sodium pyruvate, 1 mM	
Hepes, 20 mM	

'Completion' of assay medium (RPMI–C)[at]

RPMI –1640 medium	500 ml
Heat–inactivated (+56°C for 30 min) foetal calf serum (HIFCS)*	10 ml
L–Glutamine, 2 mM	
MEM non–essential amino acid, 0.1 mM	
Bovine insulin, 0.23 IU/ml	
Oxaloacetic acid, 1 mM	
Sodium pyruvate, 1 mM	
Hepes, 20 mM	

After completion, RPMI-M and RPMI-C can be stored at +4°C during three weeks.

('low–endotoxin' heat–inactivated foetal calf serum is not a crucial reagent and may be substituted by human serum AB. Also penicillin and streptomycin may be added to the RPMI-M to reduce the risk of cultures becoming contaminated)*

7.2 CELL CULTURE PROCEDURES

Starting up a cell culture [at]

Take an ampoule of frozen cells out of the liquid nitrogen container and put the ampoule on ice to thaw the cells gradually. Continue with the procedure as soon as the cells are thawed. Clean the outer surface of the ampoule with ethanol 70%. Transfer the cells to a 50 ml centrifuge tube and add 10 ml medium (+4°C). Centrifuge at 100 x g for 5 min (at +4°C). Decant the supernatant carefully and resuspend the cells in 10 ml RPMI–M (+4°C). Centrifuge at 100 x g for 5 min at (+4°C). Decant the supernatant carefully and resuspend the cells in 2 ml RPMI–M. Add 8 ml RPMI–M to a 25 cm^2 tissue culture flask and transfer the 2 ml cell suspension to the flask. Check the quality of the cells visually, using a microscope. The cells should not clump together.
Incubate the cells in a CO_2–incubator (37°C, 5% CO_2, high humidity)
(*Note: As an alternative, the cells may be thawed at 37°C and washed at RT*)

Propagation of the cells [at]

Take the culture flask from the CO_2–incubator. Estimate the number of viable cells using Trypan blue exclusion: viable cells exclude Trypan blue. Take an aliquot of 100 µl of the culture flask and add 850 µl RPMI and 50 µl 0.4% w/v Trypan blue solution. Count the number of viable cells in this solution using a haematocytometer. This

procedure is described in detail in the Sigma (Biochemical and Reagents For Life Science Research) Catalogue 2000/2001, pages 1848–9.
Centrifuge the cell suspension 100 x g for 8 min at RT. Pour off the supernatant phase, resuspend the cell pellet in 4 ml RPMI-M using a serological pipette (gently aspirate and expel several times, avoid bubbles, do not vortex).

Add a part of the cell–suspension to a flask and add new medium to the cells until a final concentration of 2 x 10^5 cells/ml. The total volume depends on the size of the culture flask (For a 25, 75, 150 and 175 cm^2 flasks use 10–15 ml, 25–30 ml, 45–55 ml and about 60 ml, respectively).
It is also valid to skip the centrifugation of the cells: Remove a part of the cell-suspension from the flask and add new medium to the cells (final concentration of 2 x 10^5 cells/ml) In general the cell suspension should be diluted between 1 in 4 up to 1 in 6.
Check the quality of the cells visually, using a microscope. The cells should not cluster. Incubate the cells in a CO_2–incubator (37°C, 5% CO_2, high humidity)

In general this procedure should be performed twice a week.

Preparation of a cell bank [at]

Centrifuge the cell culture at 100 x g for 8 min at RT (or +4°C). Decant the supernatant carefully and resuspend the cells in FBS at RT (or +4°C). Adjust the cell concentration to ≥ 4 x 10^6 cells/ml. Put the cell suspension on ice for 10 minutes.
Add dropwise an equal volume of a cold solution of FBS + 10% DMSO to the cell suspension (final cell concentration is ≥ 2 x 10^6 cells/ml with 5% DMSO).
Transfer the cell suspension to sterile and pyrogen free cryotubes (1 ml/tube) and close the tubes firmly.
Put the tubes in a well insulated polystyrene box and store the box at –70 °C or below for about 48 h. Finally transfer the tubes to a liquid nitrogen container.

7.3 ONSET OF THE TEST

Pre-incubation of cells for a test [at]

Centrifuge 30–50 ml of cell suspension at 100 x g for 8 min at room temperature. Pour off the supernatant phase, and resuspend the cells in approximately 2 ml of RPMI–C using a serological pipette (gently aspirate and expel several times, avoid bubbles, do not vortex). Count the cells and dilute the needed amount of viable cells in RPMI-C until a final concentration of 4x10^5 cells/ml. The total volume depends on the size of the culture flask and the number of cells needed for the test. (Generally 2x10^7 cells pre-incubated in 50 ml RPMI-C in a 150 cm^2 flask is enough for one 96-well assay plate). Incubate the cells during approximately 24 hours in a CO_2–incubator (37°C, 5% CO_2, high humidity)

Preparation of cells for a test [at]

Centrifuge 30–50 ml of cell suspension at 100 x g for 8 min at room temperature. Pour off the supernatant phase, and resuspend the cells in approximately 2 ml of RPMI–C using a serological pipette (gently aspirate and expel several times, avoid

bubbles, do not vortex). Count the cells and dilute the viable cells with RPMI–C to a volume that gives a concentration of 2.5 x 10^6 viable cells/ml. (Each 96–well assay plate requires about 10 ml of 2.5 x 10^6 viable cells/ml.) Prepare the solution of cells just prior to addition to the culture plate.

Preparation of samples for test[at]

Samples are tested at a dilution of 1 in 5, i.e. 50 µl of sample in a total culture volume of 250 µl. To test samples at dilutions greater than 1 in 5, pre–dilute samples before addition to the assay plate, e.g. to test a sample at a dilution of 1 in 10, pre–dilute the sample 1 in 2 with saline and add 50 µl of this diluted sample to the assay plate.

Equilibration of reagents for the test

Bring a vial of the LPS standard, the samples for assay and a bottle of RPMI–C to room temperature.

Preparation of the LPS standard curve [at]

Prepare the LPS standard curve by making serial dilutions in saline of an aliquot of the stock solution of the current IS.

Label seven tubes, A – G. Add the volumes of saline to the tubes specified in table 1, below.

Add 1.35 ml saline to an aliquot (300 IU=EU in 150 µl) of the LPS standard and vortex to make 1.5 ml of a 200 IU/ml solution of LPS = Solution S.

Table 1. Preparation of the LPS standard curve:

Tube	LPS added to tube	Saline	[LPS] in tube	→ *[LPS] in well*
A	100 µl of Solution S = 20 IU	900 µl	20 IU/ml	*Not for culture*
B	200 µl of Solution A = 4 IU	800 µl	4 IU/ml	*0.8 IU/ml*
C	500 µl of Solution B = 2 IU	500 µl	2 IU/ml	*0.4 IU/ml*
D	500 µl of Solution C = 1 IU	500 µl	1 IU/ml	*0.2 IU/ml*
E	500 µl of Solution D = 0.5 IU	500 µl	0.5 IU/ml	*0.1 IU/ml*
F	500 µl of Solution E = 0.25 IU	500 µl	0.25 IU/ml	*0.05 IU/ml*
G	None	1 ml	0 IU/ml	*0 IU/ml*

Vortex each of Solutions A – G after its preparation and then use each solution for the preparation of the subsequent dilution.

(Multiples of the above volumes may be used to generate larger volumes of Solutions A – G)

Testprocedure

Add 50 µl of LPS standards to wells as in Template 1, below.

Solution G into wells A3 – D3 (STD R0)
Solution F into wells A4 – D4 (STD R1)
Solution E into wells A5 – D5 (STD R2)
Solution D into wells A6 – D6 (STD R3)
Solution C into wells A7 – D7 (STD R4)
Solution B into wells A8 – D8 (STD R5)

(The above order of addition permits the same tip to be used for additions of all the standards.)

Add 50 µl of the test samples S1 – S14 to wells as in Template 1, see below.

Add 100 µl of RPMI–C to the wells of columns 1 – 10 as in Template 1, see below.

Gently swirl the solution of MONOMAC 6 cells to reduce settling of the cells and to distribute the cells more evenly throughout the RPMI–C solution immediately before aliquots of cells are taken. Do not vortex.

Add 100 µl of MONOMAC–6 cells to the wells of columns 1 – 10 as in Template 1, see below.

Using a pipette with a tip of wide diameter, the cells are added by row in the following sequence: A, E, B, F, C, G, D, H (see template 1, below). A repeating pipette may be used for these additions provided that the aliquots are added briskly to minimise the settling of cells.

Gently swirl the resulting cultures to mix the contents of the wells without cross–contaminating wells.

Remark: It is also valid to test the sample in a total culture volume of 200 µl. Take 40 µl of sample dilution, 110 µl medium and 50 ul of a cell solution (cell concentration 4 x 10^6 viable cells/ml)

Incubate the cultures without vibration (to allow the cells to settle) at 37°C for 16 – 24h in an atmosphere of 5% CO_2 in humidified air.

Template 1: MONOMAC 6 CELLS culture plate (example)

	1	2	3	4	5	6	7	8	9	10	11	12
A	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	void	void
B	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	void	void
C	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	void	void
D	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	void	void
E	S5	S6	S7	S8	S9	S10	S1 1	S1 2	S1 3	S1 4	void	void
F	S5	S6	S7	S8	S9	S10	S1 1	S1 2	S1 3	S1 4	void	void
G	S5	S6	S7	S8	S9	S10	S1 1	S1 2	S1 3	S1 4	void	void
H	S5	S6	S7	S8	S9	S10	S1 1	S1 2	S1 3	S1 4	void	void

Key:
S1 – S14 = test samples #1 – #14
R0 – R5 = Reference Standard Endotoxin, R0 = 0 IU/ml, R1 = 0.25 IU/ml, R2 = 0.5 IU/ml, R3 = 1 IU/ml, R4 = 2 IU/ml and R5 = 4 IU/ml (The final concentrations are: 0.05, 0.1, 0.2, 0,4 and 0.8 IU/ml).

Harvesting and storage of supernatants.

At the end of the cell culture incubation the supernatant is harvested as follows:
Transfer from each well ca. 140 µl (or more if possible) of the supernatant above the MONOMAC 6 cells to the corresponding well a new 96-wells cell culture plate. Do not disturb the cells. The use of an eight channel pipette will facilitate the transfer. Change the pipette tips between each column transfer.
Cover the plate with the supernatants with a seal. Store the plate at –70 °C (or below) until required or proceed with the detection of IL6. When the plates are stored at –20 °C, determination of IL6 should be performed within two weeks.

7.4 DETECTION OF IL–6 IN THE SUPERNATANT MEDIUM BY ELISA

Immunoreactive IL–6 in aliquots of the cell culture fluid (cell–conditioned medium) is quantified using a validated ELISA, in which the IS for IL–6 (89/548) or an IL–6 standard calibrated against the IS is used as the assay calibrant.
Two validated IL-6 ELISA's are described in the ANNEX. (The NOVARTIS ELISA for IL–6 and the Human IL–6 ELISA kit (CLB, Amsterdam, The Netherlands, code M1916) Other validated ELISAs may be substituted.

8. DATA ANALYSIS

All experiments are run with four replicates of the test compound on one plate. A standard curve using the International Standard for Endotoxin (calibrated in IU) is included, ranging from 0.25 IU/ml up to 4 IU/ml. A test is valid if the 0.5 IU/ml is significant elevated over the background (defined by mean +2SD (n-1)). The endotoxin value of each replicate is calculated from the endotoxin calibration curve of the IS for endotoxin, applying the 4-parameter logistic model and expressed as endotoxin-equivalents/ml (EU/ml). Outliers are rejected only after checking according to the Dixon's test ($p \geq 0.05$). Subsequently, the mean endotoxin value of all replicates (usually quadruplicates) of a test compound is calculated and multiplied by the dilution factor (if applicable). The mean endotoxin concentration is compared with the endotoxin limit concentration (ELC) for the test compound. Where the ELC is not specified for a product, it is calculated as described below in the prediction model.

Alternative analyses are permitted provided these are consistent with the relevant ICH guidelines.

Test validation and interference testing

To assure the precision or validity of the test method, preparatory tests are conducted to assure that the criteria for the standard curve are valid and that the test solution does not interfere in the test. The test method is validated and a test for interfering factors repeated whenever there is any change in either the test method or the test preparation that is likely to influence the result of the test.

Interference test cell culture

Spike an aliquot of the test solution of the preparation being examined with an endotoxin concentration at or near the middle of the endotoxin standard curve. Calculate the mean recovery of the added endotoxin by subtracting the mean endotoxin concentration in the test solution (if any) from that containing the added endotoxin. Calculate the spike recovery for each dilution in percent, taking the theoretical value (spike concentration e.g. 1 EU/ml) as a 100%. The test solution is considered free of interfering factors if the measured concentration of the endotoxin added to the test solution is within 50–200% of the known added endotoxin concentration after subtraction of any endotoxin detected in the solution without added endotoxin. If the test solution interferes in the test, i.e. does not give 50–200% endotoxin spike recovery, the test for interfering factors is repeated with the test solution diluted as far as is necessary to achieve 50–200% endotoxin spike recovery. The lowest dilution (highest concentration) of the product that yields an endotoxin spike recovery of 50-200% is determined. This should not exceed the maximum valid dilution (MVD) for the product. The MVD is the maximum allowable dilution of a preparation at which the endotoxin limit concentration (ELC) can be determined (MVD = ELC/limit of detection of the test).

Interference test readout system

An unknown test compound should be tested for possible interference with the IL6-ELISA itself. For this purpose, a IL6 concentration at or near the middle of the IL6 standard curve is mixed with the various dilutions of the test compound in RPMI-C

Institute for Health and Consumer Protection

(and as a control with RPMI-C alone) and subsequently tested in the IL-6 ELISA. The test compound dilutions should be comparable to those used in the cell culture test.

Prediction model

For every kind of test compound the interference in the test cell culture and the test readout system should be determined. The test compound should be tested in the lowest dilution showing no interference (endotoxin spike recovery of 50-200%) and not exceeding the maximum valid dilution for the product. The test should be performed according to the SOP and the endotoxin concentration of the test compound should be calculated as described above. The test compound is considered pyrogenic when the endotoxin concentration of the test preparation exceeds the endotoxin limit concentration (ELC) for the preparation.

Where an ELC is not specified, it is calculated as follows:
The sensitivity of rabbits to endotoxin is 5 IU/kg. So, for a product injected (i.e. tested) at 1 ml/kg, the detection limit is 5 IU endotoxin/ml/kg, giving an ELC of 5 IU endotoxin/ml, whereas for a product injected at 10 ml/kg, the detection limit is 5 IU endotoxin/10 ml/kg = 0.5 IU endotoxin/ml/kg, giving an ELC of 0.5 IU endotoxin/ml.

9. HEALTH SAFETY AND ENVIRONMENT

Human material

Human material should be treated as biologically hazardous and all work using human material is to be carried out according to laboratory safety procedures.

Cultures of human material should be treated as biologically hazardous waste and disposed of according to the laboratory safety procedures.

Bacterial endotoxin is, as its name indicates, a toxic agent and should be handled with care.

Precautions: Cover open cuts before use. Do not get in eyes, on skin, on clothing. Avoid inhaling. Keep container closed.

First Aid: In case of contact, immediately flush eyes or skin with plenty of water for at least 15 minutes. If inhaled, remove to fresh air. If not breathing, give artificial respiration, preferably mouth–to–mouth. If breathing is difficult, give oxygen.

Effects of skin absorption can include fever, headache and hypotension.

Effects of inhalation can include fever, headache and hypotension.

Effects of ingestion – adverse effects are unlikely since ingested endotoxin is rapidly detoxified.

ELISA Substrate: TMB ***(Suspected mutagen, wear gloves when handling).*** Store and use in accordance with manufacturer's instructions.

11. REFERENCES

Bleeker, W.K., de Groot, E.M., den Boer, P.J., Biessels, P.T., Aarden, L.A. and Bakker, J.C. (1994) Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol. 22, 835–40.

Dinarello, C.A., Gatti, S. and Bartfai, T. (1999) Fever: links with an ancient receptor. Curr Biol. 9, R147–50.

Dinarello, C.A., O'Connor, J.V., LoPreste, G. and Swift, R.L. (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J. Clin. Microbiol. 20, 323–329.

Duff, G.W. and Atkins, E. (1982) The detection of endotoxin by *in vitro* production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52, 323–332.

Fennrich, S., Fischer, M., Hartung, T., Lexa, P., Montag–Lessing, T., Sonntag, H.G., Weigandt, M. and Wendel. (1999) A Detection of endotoxins and other pyrogens using human whole blood. Dev Biol Stand. 101, 131–9.

Finney DJ: Statistical Method in Biological Assay, Third Edition. London, Charles Griffin and Company Ltd, 1978.

Gaines Das RE and Tydeman MS. (1982) Iterative weighted regression analysis of logit responses: A computer program for analysis of bioassays and immunoassays. Computer Programs in Biomedicine 15; 13–22.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15, 425–33.

Levin, J. and Bang, F.B. (1964) A description of cellular coagulation in the Limulus. Bull. John Hopkins Hosp. 155, 337–345.

Mascoli, C.C. and Weary, M.E. (1979a) Limulus amebocyte lysate (LAL) test for detecting pyrogens in parenteral injectable products and medical devices: advantages to manufacturers and regulatory officials. J Parenter Drug Assoc. 33, 81–95.

Mascoli, C.C. and Weary, M.E. (1979b) Applications and advantages of the Limulus amebocyte lysate (LAL) pyrogen test for parenteral injectable products. Prog Clin Biol Res. 29, 387–402.

Poole, S., Selkirk, S., Rafferty, B., Meager, A., Thorpe, R. and Gearing, A. (1989) Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa, November 1989, 17–18.

Institute for Health and Consumer Protection

Poole, S., Thorpe, R., Meager, A., Hubbard, A.R. and Gearing, A.J.H. (1988) Detection of pyrogen by cytokine release. Lancet 8577, 130.

Ray, A., Redhead, K., Selkirk, S. and Poole, S. (1990) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211–218.

Taktak, Y.S., Selkirk, S., Bristow, A.F., Carpenter, A., Ball, C., Rafferty, B. and Poole, S. (1991) Assay of pyrogens by interleukin–6 release from monocytic cell lines. J Pharm Pharmacol 43, 578–582.

11. ANNEX

NOVARTIS IL–6 ELISA:

Materials, reagents.

Nunc-immuno MaxiSorp F96)
Mouse monoclonal anti–IL–6 antibody from clone 16 (Novartis)
Horseradish peroxidase conjugated sheep polyclonal anti–IL–6 antibody (Novartis)
3,3',5,5'–Tetramethyl benzidine (e.g. Fluka Cat. No. 87748)
Acetone (reagent grade)
Ethanol (reagent grade)
Phenol (e.g. Merck Cat. No. 100206)
Potassium hydroxide (reagent grade)
Sodium hydroxide (reagent grade)
Hydrochloric acid (reagent grade)
Sodium dihydrogen phosphate (e.g. Merck Cat. No. 106346)
Disodium hydrogen phosphate (e.g. Merck Cat. No. 106580)
Tris (hydroxymethyl) aminomethane (e.g. Fluka Cat. No. 93352)
Kathon MW/WT, Christ Chemie AG, Reinach, Switzerland
Albumin from bovine serum (e.g. Fluka Cat. No. 05480)
Citric acid monohydrate (e.g. Fluka Cat. No. 27490)
Sulphuric acid (2 M H_2SO_4, reagent grade)

Preparation of buffers

Coating Buffer
Dissolve 5.0 g sodium dihydrogen phosphate and 2.9 g disodium hydrogen phosphate in 400 ml distilled water. Use 1 N NaOH to adjust the pH to 7.5, and make up to 500 ml with distilled water.
Remains stable for 6 months at 2–8°C.

Blocking Buffer
Dissolve 12.1 g Tris (hydroxymethyl)aminomethane in 400 ml distilled water.
Add 0.1 ml Kathon MW/WT. Use 4 M HCl to adjust the pH to 7.5.
Add 5.0 g BSA. Add distilled water to make up to 500 ml.
Remains stable for 6 months at 2–8°C.

Stopping Solution
Add 26.6 ml H_2SO_4 to 500 ml distilled water.

Wash Solution
Add 1 ml Tween –20 to 2000 ml of demineralised water.

Dillution Buffer
Dissolve 2.1 g Tris(hydroxymethyl)aminomethane in 400ml distilled water.
Add 0.1 ml Kathon MW/WT, 0.5 g phenol and 25 ml heat–inactivated (30 minutes at +56°C) foetal bovine serum.

Mix to dissolve the substances, then adjust the pH to 7.5 with 4 M HCl. Make up to 500 ml with distilled water.
Remains stable for at least 6 months at 2–8 C.
In the absence of the stabilisers Kathon and phenol the stability is only 1 day.

TMB Solution #
Dissolve 240 mg TMB in 5 ml acetone. Add 45 ml ethanol and 0.3 ml Perhydrol (30 % H_2O_2).
Remains stable for at least 6 months at 15–25°C when sealed and protected from light.

Substrate Buffer#
Dissolve 6.3 g citric acid monohydrate in 800 ml distilled water.
Adjust the pH to 4.1 by adding 4 M KOH.
Make up to 1000 ml with distilled water and add 0.2 ml of Kathon MW/WT.

Remains stable for about 6 months at 15–25°C.
In the absence of the Kathon the stability is only 1 day.

(# TMB solution and substrate buffer may be replaced by a TMB ready-to-use substrate system (e.g. Sigma, T8665)).

Coating of IL–6 ELISA plates

For the NOVARTIS IL–6 ELISA, dilute the coating anti–IL–6 antibody (Clone 16) with coating buffer to 2.5 μg/ml and swirl to mix, e.g. 1 mg of antibody in 400 ml of coating buffer. Add 200 μl to each well of a 96–well plate (Nunc–Immuno MaxiSorp F96). Stack the microtitre plates and allow to stand in the dark at 15–25°C for 16–24 h.

Aspirate and discard the coating solution. Wash the coated plate 3 times with demineralised water and tap out onto absorbent material, e.g. paper towel. Pipette 200 µl of blocking buffer into each well to block the residual protein–binding capacity of the coated plates. Seal the microtiter plates with adhesive film and store in a humidified atmosphere at 2–8°C (shelf life: two months).

DETECTION OF IL-6 IN THE CELL CONDITIONED MEDIUM BY ELISA

Equilibration of reagents

Bring an aliquot of the IL–6 standard and other assay reagents to room temperature before proceeding.

Preparation of IL–6 standard curve

About 30 min before the end of the tissue culture, prepare the IL–6 standard curve by making serial dilutions, in RPMI–C, of an aliquot of the stock solution of the current IS (or working STD for IL–6 calibrated against the IS).

When using the IS for IL–6, label nine tubes H – P. Add the specified volumes of RPMI–C to the tubes – see table 2 below.

Add 180 µl RPMI–C to an aliquot (20 ng in 20 µl) of the IL–6 standard and vortex to make 200 µl of a 100 ng/ml solution of IL–6 = Solution G.

Table 2. Preparation of the IL–6 standard curve

Tube	IL–6 added	RPMI–C	[IL–6] in tube	→ *[IL–6] in well*
H	100 µl of Solution G = 10 ng	900 µl	10 ng/ml	*Not for ELISA*
I	800 µl of Solution H = 8 ng	1.2 ml	4 ng/ml	*4000 pg/ml*
J	1 ml of Solution I = 4 ng	1 ml	2 ng/ml	*2000 pg/ml*
K	1 ml of Solution J = 2 ng	1 ml	1 ng/ml	*1000 pg/ml*
L	1 ml of Solution K = 1ng	1 ml	500 pg/ml	*500 pg/ml*
M	1 ml of Solution L = 500 pg	1 ml	250 pg/ml	*250 pg/ml*
N	1 ml of Solution M = 250 pg	1 ml	125 pg/ml	*125 pg/ml*
O	1 ml of solution N = 125 pg	1 ml	62.5 pg/ml	*62.5 pg/ml*
P	None	2 ml	0 pg/ml	*0 pg/ml*

Vortex each of Solutions H – P after its preparation and then use each solution for the preparation of the subsequent dilution.

(Multiples of the above volumes may be used to generate larger volumes of Solutions H – P and an IL–6 standard calibrated against the IS for IL–6 may be substituted for the IS.)

Store Solutions H – P at 2– 8°C until required.

Addition of samples

Immediately prior to adding standards and samples, empty the blocking buffer from the ELISA plate(s) and expel any remaining fluid by tapping the inverted plate onto absorbent material, e.g. paper towel.
Make the supernatant more homogeneous by aspirating and expelling three times before transferring the test- amount of supernatant from each of the wells of columns 1 – 10 of the tissue culture plate into the corresponding wells on the ELISA plate – see Template 1, above and Template 2, below. (The wells in columns 11 and 12 are for the cytokine standard curve – see below). The use of an eight channel pipette will facilitate the mixing and transfer. Change the pipette tips between each column transfer.

Addition of standards

Add 50 µl of IL–6 standards to the wells in columns 11 and 12, as shown in Template 2, below. Start at the lowest concentration to permit using the same tip for additions of all the standards.

Template 2: ELISA plate

	1	2	3	4	5	6	7	8	9	10	11	12
A	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	0	0
B	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	62.5	62.5
C	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	125	125
D	S1	S2	R0 0	R1 0.25	R2 0.5	R3 1	R4 2	R5 4	S3	S4	250	250
E	S5	S6	S7	S8	S9	S10	S1 1	S1 2	S1 3	S1 4	500	500
F	S5	S6	S7	S8	S9	S10	S1 1	S1 2	S1 3	S1 4	1000	1000
G	S5	S6	S7	S8	S9	S10	S1 1	S1 2	S1 3	S1 4	2000	2000
H	S5	S6	S7	S8	S9	S10	S1 1	S1 2	S1 3	S1 4	4000	4000

Key:
S1 – S14 = test samples #1 – #14
R0 – R5 = Reference Standard Endotoxin, R0 = 0 IU/ml, R1 = 0.25 IU/ml, R2 = 0.5 IU/ml, R3 = 1 IU/ml, R4 = 2 IU/ml and R5 = 4 IU/ml (The final concentrations are: 0.05, 0.1, 0.2, 0,4 and 0.8 IU/ml).
Values 1 – 4000 in columns 11 and 12 are concentrations in pg/ml of the IS for IL–6 (ampoule code 94/580, 1 μg/ampoule, 1 pg = 0.1 IU).

Addition of (2nd) antibody–HRP conjugate (POD)

Add 200 µl of detection antibody POD (horseradish peroxidase conjugated to sheep anti–IL–6 antibodies: stable for at least 6 months at 2–8°C) pre–diluted with dilution buffer (usually 1/200 to 1/500, as determined in optimisation experiments) to each well, seal the plates with adhesive film, and allow to stand for 2–3 hours at 20–25°C. (100 ml of diluted POD is sufficient for 4 ELISA plates.)

After incubation, wash plate three times with about 250 µl per well wash solution and then three times with demineralised water. Empty plate and expel any remaining fluid by tapping the inverted plate onto absorbent material, e.g. paper towel.

Addition of substrate solution and reading of optical densities

Prepare the substrate solution shortly before use. Transfer 90ml of substrate buffer to a plastic bottle, add 4.5 ml of TMB solution and mix.

Pipette 200µl of substrate solution into each well. After 10–15 minutes, stop the enzyme reaction by adding 50 µl/well of stopping solution. Wipe the back of the microtitre plates with a clean tissue, then measure the absorbance at 450 nm in an ELISA plate reader using a 540–590nm corrective filter. Subtract the values of the measurement with the corrective filter from values measured with the 450 nm filter.

IL-6 ELISA kit (CLB, Amsterdam, The Netherlands, code M1916)

The reagents provided in the ELISA kit used are:

- Coating antibody, 100-fold concentrated
- Blocking reagent, 50-fold concentrated
- IL-6 standard, 4500 pg/ml (calibrated against the WHO International Standard)
- Biotinylated IL-6 antibody, 100-fold concentrated
- Streptavidin-HRP conjugate, 10,000-fold concentrated
- Dilution buffer, 5-fold concentrated
- Microtiter plates + lid (Nunc-Immuno Maxisorp F96)
- Plate seals

Additional materials needed are:
Carbonatebuffer (pH 9.6)
Ethanol 96%
Tween 20
Phosphate buffered saline
Sulphuric acid (2 M H_2SO_4, reagent grade)
3,3, 5,5 Tetramethylbenzidine
Peroxide (30 % H_2O_2)
sulphuric acid (2M)
sodium-acetate (1.1 M, pH 5.5)
Instead of the last four items, a TMB ready-to-use substrate system can be applied (e.g. Sigma, T-8665)

Preparation of reagent solutions used in the IL-6 ELISA
The following solutions are prepared prior to each test:

Washing buffer
Add 50 µl Tween 20 (or 500 µl 10% Tween 20) to 1000 ml PBS.

Dilution buffer
The ELISA kit contains one bottle with 5-fold concentrated dilution buffer. Calculate the quantity of dilution buffer required and prepare a working-strength dilution by diluting the concentrated buffer 1 in 5 in distilled water.

Substrate
Dissolve 6 mg TMB in 1 ml 96% ethanol. Add 1.2 ml sodium-acetate (1.1 M, pH 5.5) and 0.2 ml TMB solution to 10.8 ml of distilled water. Add 2.4 µl H_2O_2 directly prior to use. (alternatively: an equivalent TMB substrate system can be applied)

Performing the ELISA
Centrifuge all vials (except the blocking reagent) before use (1 min. at 3000 g). Preferably an ELISA-plate shaker is used during the incubations (except during the coating of the plate))

Coating of the plate

Dilute 120 µl of the coating antibody with 12 ml of carbonate buffer (= 1/100). Add 100 µl of this dilution to each well of the microtiterplate and incubate overnight at RT.

Wash the plate four times using washing buffer (platewasher).
.

Blocking of the plate
Dilute 500 µl blocking reagent with 25 ml PBS (= 1/50). Add 200 µl of this dilution to each well of the microtiterplate and incubate during 1 hour at RT. In the meantime bring the storage plate with the supernatant to RT.
Wash the plate four times using washing buffer.

Preparation of the IL6 Standard (provided with the IL6 ELISA kit)
Dilute in polystyrene tubes.

Tube	IL–6 added	dilution buffer	[IL–6] in tube	→ [IL–6] in well
1	50 µl IL6 Standard (4500pg/ml)	450 µl	450 pg/ml	450 pg/ml
2	200 µl of solution 1 (450 pg/ml)	400 µl	150 pg/ml	150 pg/ml
3	200 µl of solution 2 (150 pg/ml)	400 µl	50 pg/ml	50 pg/ml
4	200 µl of solution 3 (50 pg/ml)	400 µl	16.7 pg/ml	16.7 pg/ml
5	200 µl of solution 4 (16.7 pg/ml)	400 µl	5.6 pg/ml	5.6 pg/ml
6	200 µl of solution 5 (5.6 pg/ml)	400 µl	1.9 pg/ml	1.9 pg/ml
7	200 µl of solution 6 (1.9 pg/ml)	400 µl	0.6 pg/ml	0.6 pg/ml
8	None	400 µl	0 pg/ml	0 pg/ml

Addition of the samples (supernatants)
See: -template 2 ELISA-plate (example)

Homogenise the supernatant in the storage plate before transferring the test-amount of supernatant from the storage plate to the ELISA. Add 100 µl of each dilution of the IL-6 standard (IL6) to the assigned wells of column 11 and 12. Add 80 µl dilution buffer and 20 µl of the sample (sample = S) to the assigned wells of the microtiterplate. As a controle for the performance of the ELISA, one of the sample can be replaced with dilution buffer. Incubate for 1 hour at RT. (Plate shaker)

Wash the plate four times using washing buffer.

Conjugate 1: Biotinylated IL-6 antibody
Add 120 µl biotinylated IL-6 antibody to 12 ml dilution buffer (= 1/100).
Add 100 µl of the conjugate dilution to each well and incubate for 1 hour at RT. (Plate shaker)

Wash the plate four times using washing buffer.

Conjugate 2: Streptavidine-Peroxidase
Add 3 µl streptavidine-HRP conjugate to 30 ml dilution buffer (= 1/10,000).
Add 100 µl of the conjugate dilution to each well and incubate for half an hour at RT. (Plate shaker)

Wash the plate four times using washing buffer.

Substrate solution
Add 100 µl of substrate solution to each well and incubate 15 minutes at RT

Stop solution
Add 100 µl of sulphuric acid (2M) to each well. Measure the absorbance at 450 nm with an ELISA plate reader.

[This Page Intentionally Left Blank]

Validation of Biomedical Testing Methods

In vitro pyrogen test using MONOMAC 6 CELLS

Standard Operating Procedure

Only the responsible of the GLP/QA Unit is allowed to make copies of this document. Extra examples can be obtained at the GLP/QA Unit. Quality Documents are only valid if they are signed by the responsible of the GLP/QA Unit in and provided with a blue copy number	Copy number

Sop-MM6v08

JOINT RESEARCH CENTRE
EUROPEAN COMMISSION
Institute for Health and Consumer Protection
ECVAM

Standard Operating procedure

In vitro pyrogen test using MONOMAC 6 TEST

Drafted by:	Name	Stephen Poole
	Date	26-09-01
	Signature	

Reviewed by:	Name	Ria Nibbeling
	Date	28-02-02
	Signature	

Reviewed by:	Name	Ria Nibbeling
	Date	23-04-02
	Signature	

Reviewed by:	Name	Esther van Tienhoven
	Date	01-10-02
	Signature	

*Owner/Trainer: Signature: Date:

PAGE OF CHANGES

Date of change/ Date of draft:	Version– number:	Changed page(s):	Summary of the change(s):	Changed by/Sign.:
280202	2		Overall revision	RN
140402	3	12	Pre-incubation of cells	RN
230402	4	15	Chapter 8	RN
011002	5		Overall revision	EvT
09/10/2002	6	14	Table 1 update	EvT
09/10/2002	6	15	Preperation of Negative and Positive Product Controls updated	EvT
29/10/2002	7	15	Sample preparation [at] 0.5 ml of the respective undiluted drug will be spiked with 23.3 µl corresponding blinded spike solution (see table 2). Then dilute the drug to its MVD according to Table 3.	
29/10/2002	7	16	*Modified Table 2. Sample preparation*	
			Modified Table 3. Dilution of the substances	
30/10/2002	8	14	*Modified Table 3. Dilution of the substances*	

TABLE OF CONTENTS Page No.

1. INTRODUCTION ***5***
3. SCOPE / LIMITATIONS ***6***
4. METHOD OUTLINE ***6***
5. DEFINITIONS / ABBREVIATIONS ***7***
6. MATERIALS ***9***
6.1. Cell line 9
6.2. Technical equipment 9
6.3. Other materials 9
6.4 Chemicals and culture media 9
7. METHODS ***11***
7.1. Steps prior to cell-culture 11
Preparation of aliquots of the LPS (endotoxin) standard (STD)[at] 11
Preparation of medium[at] 11
Completion of maintenance culture medium (RPMI–M)[at] 11
Completion of assay medium (RPMI–C)[at] 11
7.2 Cell culture procedures 12
Starting up a cell culture [at] 12
Propagation of the cells [at] 12
Preparation of a cell bank [at] 12
7.3 Onset of the test 12
Pre-incubation of cells for a test [at] 13
Preparation of cells for a test [at] 13
Equilibration of reagents for the test 13
Preparation of the LPS standard curve [at] 13
Preparation of samples for test[at] 14
Test procedure 15
Template 1: MONOMAC 6 CELLS culture plate (example) 15
Harvesting and storage of supernatants 16
7.4 Detection of IL–6 in the supernatant medium by ELISA 16
8. DATA ANALYSIS ***17***
8.1 Assay acceptance criteria 17
8.2 Prediction model 17
9. HEALTH SAFETY AND ENVIRONMENT ***18***
10. REFERENCES ***19***
11. ANNEX ***21***

THIS SOP WAS AMENDED FOR THE VALIDATION PHASE ONLY. IT DOES THEREFORE ONLY REPLACE THE PREVIOUS VERSION FOR THIS SERIES OF EXPERIMENTS.

1. INTRODUCTION

Parenteral pharmaceutical products must be shown to be free from pyrogenic (fever–inducing) contamination. While a pyrogen may in general be defined as any substance that causes fever, the pyrogens that almost invariably contaminate parenteral pharmaceuticals are bacterial endotoxins (lipopolysaccharides, LPS) from Gram–negative bacteria (Mascoli and Weary, 1979a, 1979b). There are two Pharmacopoeial tests for pyrogenic contamination: the rabbit pyrogen test and the Limulus amoebocyte lysate (LAL) test. The rabbit pyrogen test, which detects LPS and other pyrogens, involves measuring the rise in body temperature evoked in rabbits by the intravenous injection of a sterile solution of the substance to be examined. In contrast, the LAL test detects only LPS: it is described in Pharmacopoeias as the bacterial endotoxins test (BET). The principle of the LAL–test is that LPS causes extracellular coagulation of the blood (haemolymph) of the horseshoe crab, *Limulus polyphemus.* (Levin & Bang, 1964). Although the LAL test is gradually superseding the rabbit pyrogen test, hundreds of thousands of rabbit pyrogen tests are still carried out each year around the world, largely on products which cannot, for one reason or another, be tested in the LAL test. While proving generally reliable, both the rabbit pyrogen test and LAL test have shortcomings. The rabbit pyrogen test uses experimental animals, is costly and is not quantitative. The LAL test gives false negatives with certain products, can overestimate the pyrogen content of other products and does not detect pyrogens other than bacterial endotoxin (LPS), such as Gram–positive exotoxins, viruses and fungi (Dinarello et al., 1984; Poole et al., 1988; Ray et al., 1990; Taktak et al., 1991; Fennrich et al., 1999).

The basis of the rabbit pyrogen test is the *in vivo* stimulation by exogenous pyrogens (usually LPS) of rabbit peripheral blood monocytes to produce the endogenous pyrogens that cause fever. The endogenous pyrogens are pyrogenic cytokines such as tumour necrosis factorα (TNFα), interleukin–1 (IL–1α and IL–1β, two separate gene products), IL–6 and IL–8 (Dinarello et al., 1999). In view of the shortcomings of the rabbit pyrogen test and the LAL test, in vitro pyrogen tests that utilise the exquisite sensitivity to exogenous pyrogen of monocytes have been proposed. In such tests, products are incubated with human peripheral blood monocytes (or mononuclear cells, PBMNC CELLS, or leukocytes) and the conditioned media assayed for pyrogenic cytokines (Duff & Atkins, 1982; Dinarello et al., 1984; Poole et al., 1988; Poole, 1989; Hansen and Christensen, 1990; Taktak et al., 1991; Bleeker et al., 1994).

The isolation of monocytes/leukocytes from whole blood is labour–intensive and time–consuming, technically sophisticated, requires expensive reagents and does not guarantee the isolation of cells in a non–activated state. This prompted the evaluation of various cell lines which retain monocytic characteristics, including the capacity to synthesise and secrete pyrogenic cytokines (Taktak et al., 1991). Taktak *et al.* measured LPS–induced IL–1β and IL–6 release from MONO MAC 6 cells and THP–1 cells and concluded that IL–6 release by MONO MAC 6 cells was the most appropriate readout for an *in vitro* pyrogen test ('monocyte test') because immunoreactive IL–6, unlike immunoreactive IL–1 and TNFα, is secreted entirely into the cell–conditioned medium in large quantities, permitting its complete estimation. The test was applied to three batches of therapeutic human serum albumin (HSA) that had caused adverse reactions in recipients. The MONO MAC 6/IL–6 test detected pyrogenic contamination in the HSA that had not been detected in the rabbit pyrogen test and the LAL test.

2. PURPOSE

To develop an *in vitro* pyrogen test that will serve as a replacement for the rabbit pyrogen test.

3. SCOPE / LIMITATIONS

The method described below is for the evaluation of an in vitro MONOMAC 6/IL–6 release test. It is not a 'finalised' test system for the testing of medicinal products. The method may be applied only to preparations that have been validated with the method.

4. METHOD OUTLINE

MONOMAC 6 cells are stimulated for 16–24h with standard endotoxin (LPS) and preparations under test. Following this stimulation, the concentration of IL–6 in the cell–conditioned medium is quantified using a specific ELISA (which is calibrated in terms of the appropriate international standard). The construction of a dose–response curve for endotoxin standard versus OD-value of IL–6 ELISA permits the estimation of the pyrogenic contamination of the preparations under test. The contamination is measured in endotoxin–equivalent units.

5. DEFINITIONS / ABBREVIATIONS

µg	microgram
µl	microlitre
Ab	antibody
BSA	bovine serum albumin
CO_2	Carbon dioxide
°C	degrees Celsius (Centigrade)
DMSO	Dimethylsulfoxide
D–R	dose–response
E. coli	Escherichia coli
ELISA	Enzyme–linked immunosorbent assay
EP	European Pharmacopoeia
EU	endotoxin units
FDA	Food and Drug Administration (USA)
g	gram
h	hour
HIFCS	heat–inactivated (+56°C for 30 min) foetal calf serum
H_2O_2	hydrogen peroxide
H_2SO_4	sulphuric acid
HSA	Human Serum Albumin
IL	interleukin
IS	international standard
IU	international unit
l	litre
KOH	potassium hydroxide
LAL	limulus amoebocyte lysate
LPS	lipopolysaccharide
M	molar
MAb	monoclonal antibody
mg	milligram
min	minute
ml	millilitre
mM	millimolar
mvd	maximum valid dilution
NaCl	sodium chloride
NaOH	sodium hydroxide
$NaHCO_3$	sodium hydrogen carbonate
NaH_2PO_4	sodium di–hydrogen phosphate
Na_2HPO_4	di–sodium hydrogen orthophosphate
No.	number
nm	nanometre
OD	optical density
PBS	Dulbecco's phosphate buffered saline
PC	Personal Computer
PF	pyrogen–free (items purchased as sterile and pyrogen–free or baked at 250°C for 30–60 min.)
POD	horseradish peroxidase conjugate
R	endotoxin standard
rpm	rounds per minute
RPMI	RPMI 1640 cell culture medium
RSE	Reference Standard Endotoxin
RT	room temperature
TMB	Tetramethyl benzidine
S	test sample

STD	standard
UNK	unknown
USP	United States Pharmacopoeia
x g	x gravity

6. MATERIALS

6.1. Cell line

The cell line used is human monocytic cell line MonoMac-6 (obtained from Prof. H.W.L. Ziegler-Heitbrock (Institute for Immunology, University of Munich, Munich, Germany) A Master Cell Bank and a Working Cell Bank was established at the NIBSC. Cells can be obtained from this laboratory.

6.2. Technical equipment

Incubator (37°C ± 2 °C, 5% ± 0.5 % CO_2, humidified)
Inverted microscope
Haematocytometer
Laminar flow clean bench (Class II)
Laboratory centrifuge (refrigerated) suitable for 50 ml tubes
Laboratory centrifuge suitable for 1 ml vials
Water bath (adjustable to 37 °C, 56°C and 70°C)
pH meter
ELISA-reader
Platewasher
Pipettes adjustable to 2-20 µl, 50-200 µl, 20-100 µl or 200-1000 µl (e.g. Gilson)
12-channel or 8-channel pipette
Vortex mixer

6.3. Other materials

All materials must be sterile and pyrogen free.

Tissue culture flasks, 25 cm^2, 75 cm^2 and 150 cm^2 (e.g. Costar)
Centrifuge tubes, 15 ml and 50 ml (e.g.Greiner or Falcon)
pyrogen free tips (e.g. Greiner)
96-wells tissue culture plates with lid (e.g.Costar)
polystyrene tubes (e.g.Greiner or Falcon)
plate sealers, non toxic (e.g.Dynatech Laboratories)
cryotubes, 2 ml (e.g. Nunc)
serological pipettes (5ml,10ml, 25ml, e.g. Beckton Dickinson Labware)
0.22 µm sterile filters (MilliPak 60, Millipore)

6.4 Chemicals and culture media

Sterile, pyrogen–free phosphate buffered saline (e.g. Life Technologies)
Hydrochloric acid, 0.1M, sterile filtered (Sigma, H–9892)
Pyrogen free sodiumchloride 0.9% (i.e. saline)(e.g. NPBI)
Water for injections (e.g. NPBI)
Trypan blue stain (e.g Sigma)

RPMI 1640 (e.g. Lifetechnologies, Gibco BRL, code 31870–025)
Hepes (e.g. 1M, Gibco)
Insulin (Sigma, code 1–4011)

L–Glutamine (e.g. 200 mM , Gibco)
Oxaloacetic acid (e.g. Sigma)
Sodium pyruvate (e.g. 100 mM, Gibco)
MEM non–essential amino acid solution (e.g. Gibco)
Fetal Bovine Serum (FBS) (e.g. Myoclone Super Plus FBS, Gibco, code 16000-036).
Note: The endotoxin content must be < 0.1 ng/ml, checked by LAL, after heating for 30 min. at 70°C.
Dimethyl sulphoxide (e.g. Merck)
Penicillin/Streptomycin, (10,000 IU/ml penicillin, 10 mg/ml streptomycin (e.g. Gibco, code 1514-0114))
International Standard for Endotoxin, 10,000 IU per vial (NIBSC, code 94/580)
(HSA, 1% (a dilution in sterile PF saline of clinical grade HSA, 4.5%))
(Human serum AB (e.g. Sigma))

All other consumables are purchased as sterile and pyrogen–free and other reagents are pro analysis grade.

7. METHODS

7.1. Steps prior to cell-culture

Steps marked ([at]) are carried out in a Class 2 laminar flow sterile cabinet, using aseptic technique and reagents and consumables that are sterile and pyrogen–free.

Preparation of aliquots of the LPS (endotoxin) standard (STD)[at]

Make aliquots of the LPS STD: take a vial of the current IS for endotoxin (vial code 94/580, 10000 IU = EU/vial, infinite shelf life when stored at –20°C or below), and reconstitute the contents of the vial with 5 ml pyrogen free water and vortex for 30 min. This gives a 2,000 IU/ml stock solution of LPS. The stock solution may be kept at 2–8°C for up to 14 days. Alternatively, aliquot LPS standard into labelled cryotubes, freeze them upright and store them at –20°C or below (shelf life = 12 months) The following primary standards (Reference Standard Endotoxins) are identical to the IS and may be substituted for it: EC6 (USP), Lot G (FDA), BRP3 (EP). Alternatively, a working standard (control standard endotoxin) calibrated against one of these primary standards may be substituted.

Preparation of medium[at]

Prepare and store the RPMI as described by the supplier. Adjust the pH and sterile filter (0.22 μm) the medium if required.
Test a sample of the medium (from one bottle) for LPS contamination in an LAL test according to the current SOP's for LAL testing. Use the batch of medium only if the level of contamination is not greater than 0.06 IU/ml.

Completion of maintenance culture medium (RPMI–M)[at]

RPMI 1640 medium		500 ml
Heat–inactivated (+56°C for 30 min) foetal calf serum (HIFCS)		50 ml
L–Glutamine, 2 mM		
MEM non–essential amino acid, 0.1 mM		
Bovine insulin, 0.23 IU/ml		
Oxaloacetic acid, 1 mM		
Sodium pyruvate, 1 mM		
Hepes, 20 mM		

Completion of assay medium (RPMI–C)[at]

RPMI –1640 medium		500 ml
Heat–inactivated (+56°C for 30 min) foetal calf serum (HIFCS)	*	10 ml
L–Glutamine, 2 mM		
MEM non–essential amino acid, 0.1 mM		
Bovine insulin, 0.23 IU/ml		
Oxaloacetic acid, 1 mM		
Sodium pyruvate, 1 mM		
Hepes, 20 mM		

After completion, RPMI-M and RPMI-C can be stored at +4°C during three weeks.

('low–endotoxin' heat–inactivated foetal calf serum is not a crucial reagent and may be substituted by human serum AB. Also penicillin and streptomycin may be added to the RPMI-M to reduce the risk of cultures becoming contaminated)*

7.2 Cell culture procedures

Starting up a cell culture [at]

Take an ampoule of frozen cells out of the liquid nitrogen container and put the ampoule on ice to thaw the cells gradually. Continue with the procedure as soon as the cells are thawed. Clean the outer surface of the ampoule with ethanol 70%. Transfer the cells to a 50 ml centrifuge tube and add 10 ml medium (+4°C). Centrifuge at 100 x g for 5 min (at +4°C). Decant the supernatant carefully and resuspend the cells in 10 ml RPMI–M (+4°C). Centrifuge at 100 x g for 5 min at (+4°C). Decant the supernatant carefully and resuspend the cells in 2 ml RPMI–M. Add 8 ml RPMI–M to a 25 cm^2 tissue culture flask and transfer the 2 ml cell suspension to the flask. Check the quality of the cells visually, using a microscope. The cells should not clump together.
Incubate the cells in a CO_2–incubator (37°C, 5% CO_2, high humidity)
(*Note: As an alternative, the cells may be thawed at 37°C and washed at RT*)

Propagation of the cells [at]

Take the culture flask from the CO_2–incubator. Estimate the number of viable cells using Trypan blue exclusion: viable cells exclude Trypan blue. Take an aliquot of 100 µl of the culture flask and add 850 µl RPMI and 50 µl 0.4% w/v Trypan blue solution. Count the number of viable cells in this solution using a haematocytometer. This procedure is described in detail in the Sigma (Biochemical and Reagents For Life Science Research) Catalogue 2000/2001, pages 1848–9.
Centrifuge the cell suspension 100 x g for 8 min at RT. Pour off the supernatant phase, resuspend the cell pellet in 4 ml RPMI-M using a serological pipette (gently aspirate and expel several times, avoid bubbles, do not vortex).

Add a part of the cell–suspension to a flask and add new medium to the cells until a final concentration of 2 x 10^5 cells/ml. The total volume depends on the size of the culture flask (For a 25, 75, 150 and 175 cm^2 flasks use 10–15 ml, 25–30 ml, 45–55 ml and about 60 ml, respectively).
It is also valid to skip the centrifugation of the cells: Remove a part of the cell-suspension from the flask and add new medium to the cells (final concentration of 2 x 10^5 cells/ml) In general the cell suspension should be diluted between 1 in 4 up to 1 in 6.
Check the quality of the cells visually, using a microscope. The cells should not cluster. Incubate the cells in a CO_2–incubator (37°C, 5% CO_2, high humidity)

In general this procedure should be performed twice a week.

Preparation of a cell bank [at]

Centrifuge the cell culture at 100 x g for 8 min at RT (or +4°C). Decant the supernatant carefully and resuspend the cells in FBS at RT (or +4°C). Adjust the cell concentration to ≥ 4 x 10^6 cells/ml. Put the cell suspension on ice for 10 minutes.
Add dropwise an equal volume of a cold solution of FBS + 10% DMSO to the cell suspension (final cell concentration is ≥ 2 x 10^6 cells/ml with 5% DMSO).
Transfer the cell suspension to sterile and pyrogen free cryotubes (1 ml/tube) and close the tubes firmly.
Put the tubes in a well insulated polystyrene box and store the box at –70 °C or below for about 48 h. Finally transfer the tubes to a liquid nitrogen container.

7.3 Onset of the test

Pre-incubation of cells for a test [at]

Centrifuge 30–50 ml of cell suspension at 100 x g for 8 min at room temperature. Pour off the supernatant phase, and resuspend the cells in approximately 2 ml of RPMI–C using a serological pipette (gently aspirate and expel several times, avoid bubbles, do not vortex). Count the cells and dilute the needed amount of viable cells in RPMI-C until a final concentration of $4x10^5$ cells/ml. The total volume depends on the size of the culture flask and the number of cells needed for the test. (Generally $2x10^7$ cells pre-incubated in 50 ml RPMI-C in a 150 cm^2 flask is enough for one 96-well assay plate). Incubate the cells during approximately 24 hours in a CO_2–incubator (37°C, 5% CO_2, high humidity)

Preparation of cells for a test [at]

Centrifuge 30–50 ml of cell suspension at 100 x g for 8 min at room temperature. Pour off the supernatant phase, and resuspend the cells in approximately 2 ml of RPMI–C using a serological pipette (gently aspirate and expel several times, avoid bubbles, do not vortex). Count the cells and dilute the viable cells with RPMI–C to a volume that gives a concentration of 2.5 x 10^6 viable cells/ml. (Each 96–well assay plate requires about 10 ml of 2.5 x 10^6 viable cells/ml.) Prepare the solution of cells just prior to addition to the culture plate.

Equilibration of reagents for the test

Bring a vial of the LPS standard, the samples for assay and a bottle of RPMI–C to room temperature.

Preparation of the LPS standard curve [at]

Prepare the LPS standard curve by making serial dilutions in saline of an aliquot of the stock solution of the current IS.

Label seven tubes: A – G. Add the volumes of saline to the tubes specified in table 1, below.

Add 1.35 ml saline to an aliquot (300 IU=EU in 150 µl) of the LPS standard and vortex extensively to make 1.5 ml of a 200 IU/ml solution of LPS = Solution S. Then prepare as indicated in Table 1.

Table 1. Preparation of the LPS standard curve:

Tube	LPS added to tube	Saline	[LPS] in tube	→ *[LPS] in well*	
A	100 µl of Solution S = 20 IU	900 µl	20 IU/ml	*Not for culture*	
B	200 µl of Solution A = 4 IU	800 µl	4 IU/ml	*Not for culture*	
C	500 µl of Solution B = 2 IU	500 µl	2 IU/ml	*0.4 IU/ml*	R5
D	500 µl of Solution C = 1 IU	500 µl	1 IU/ml	*0.2 IU/ml*	R4
E	500 µl of Solution D = 0.5 IU	500 µl	0.5 IU/ml	*0.1 IU/ml*	R3
F	500 µl of Solution E = 0.25 IU	500 µl	0.25 IU/ml	*0.05 IU/ml*	R2
G	500 µl of Solution F = 0.125 IU	500 µl	0.125 IU/ml	*0.025 IU/ml*	R1
H	None	1 ml	0 IU/ml	*0 IU/ml*	R0

Thoroughly vortex each of Solutions A – G after its preparation and then use each solution for the preparation of the subsequent dilution.

(Multiples of the above volumes may be used to generate larger volumes of Solutions A – G)

Preparation of samples for test [at]

Sample preparation [at]
0.5 ml of the respective undiluted drug will be spiked with 23.3 µl corresponding blinded spike solution (see table 2). Then dilute the drug to its MVD according to Table 3.

Preparation of the Negative Product Control (NPC) [at]
Dilute each testproduct to its MVD to a total volume of 0.5 ml. Spike with 25 µl of pyrogen free saline (see table 2).

Preparation of the Positive Product Control (PPC) [at]
The Positive Product Control (PPC) is used for the prediction model of the validation study as demarcation value to discriminate between positive and negative testproducts.
Dilute each testproduct to its MVD to a total volume of 0.5 ml. Spike with 25 µl of the unblinded PPC-LPS spike solution (1:21 dilution of the spike, final concentration 50 pg/ml) (see table 2).

Preparation of the Positive assay Control (PC) [at]
Take 0.5 ml of saline. Spike with 25 µl of the unblinded PPC-LPS spike solution (1:21 dilution of the spike, final concentration 50 pg/ml)

Preparation of the Negative assay Control (NC) [at]
Take 0.5 ml of saline.

Table 2. Sample preparation

unblinded			*blinded*				
dilution of drug up to **MVD**			Spiking of **undiluted** drug: 0.5 ml each				
diluted drug	**NPC**	**PPC**	+ 23.3 µl of Spike 1	+ 23.3 µl of Spike 2	+ 23.3 µl of Spike 3	+ 23.3 µl of Spike 4	+ 23.3 µl of Spike 5
0.5 ml	+ 25 µl saline	+ 25 µl PPC-LPS-spike *					
	test	(final conc. = 50 pg/ml) test	**dilution** to **MVD** test	test	test	test	test

* *PPC-LPS-spike contains 1050 pg/ml = 21fold 50 pg/ml*
NPC = Negative Product Control, PPC = Positive Product Control,
MVD = Maximal Valid Dilution

Table 3. Dilution of the substances

	Substance	*MVD (ELC/0.5 EU/ml)*	*_l of substance*	*_l of saline*
1	Glucose 5%	70	40	2760
2	EtOH 13%	35	80	2720

3	MCP	350	8	2792
4	Orasthin	700	4	2796
5	Binotal	140	20	2780
6	Fenistil	175	16	2784
7	Sostril	140	20	2780
8	Beloc	140	20	2780
9	Drug A	35	80	2720
10	Drug B	70	40	2760

Each substance has to be vortexed for about 5 seconds immediately before use.

Test procedure

Add 50 µl of LPS standards to wells as in Template 1, below.

Solution G into wells A3 – D3 (STD R0)
Solution F into wells A4 – D4 (STD R1)
Solution E into wells A5 – D5 (STD R2)
Solution D into wells A6 – D6 (STD R3)
Solution C into wells A7 – D7 (STD R4)
Solution B into wells A8 – D8 (STD R5)

(The above order of addition permits the same tip to be used for additions of all the standards.)

Add 50 µl of the samples to wells as in Template 1, see below.
Add 50 µl of the PPC to wells as in Template 1, see below.
Add 50 µl of the NPC to wells as in Template 1, see below.
Add 50 µl of the PC to wells as in Template 1, see below.
Add 50 µl of the NC to wells as in Template 1, see below.

Add 100 µl of RPMI–C to the wells of columns 1 – 12 as in Template 1, see below.

Gently swirl the solution of MONOMAC 6 cells to reduce settling of the cells and to distribute the cells more evenly throughout the RPMI–C solution immediately before aliquots of cells are taken. Do not vortex.

Add 100 µl of MONOMAC–6 cells to the wells of columns 1 – 12 as in Template 1, see below.

Template 1: MONOMAC 6 CELLS culture plate (example)

	1	2	3	4	5	6	7	8	9	10	11	12
A	D1-NPC	D1-PPC	R0	R1	R2	R3	R4	R5	D2-2	D2-4	NC	PC
B	D1-NPC	D1-PPC	R0	R1	R2	R3	R4	R5	D2-2	D2-4	NC	PC
C	D1-NPC	D1-PPC	R0	R1	R2	R3	R4	R5	D2-2	D2-4	NC	PC
D	D1-NPC	D1-PPC	R0	R1	R2	R3	R4	R5	D2-2	D2-4	NC	PC

E	D2-NPC	D2-PPC	D1-1	D1-2	D1-3	D1-4	D1-5	D2-1	D2-3	D2-5	void	void
F	D2-NPC	D2-PPC	D1-1	D1-2	D1-3	D1-4	D1-5	D2-1	D2-3	D2-5	void	void
G	D2-NPC	D2-PPC	D1-1	D1-2	D1-3	D1-4	D1-5	D2-1	D2-3	D2-5	void	void
H	D2-NPC	D2-PPC	D1-1	D1-2	D1-3	D1-4	D1-5	D2-1	D2-3	D2-5	void	void

D = testproduct (drug)
NPC= negative product control
PPC= positive product control
*R0 – R5 = Reference Standard Endotoxin, R0 = 0 IU/ml, R1 = 0.125 IU/ml, R2 = 0.25 IU/ml, R3 = 0.5 IU/ml, R4 = 1 IU/ml and R5 = 2 IU/ml (The final concentrations are: 0.*05, 0.1, 0.2, 0,4 and 0.8 IU/ml).
NC= negative control
PC= positive control
Using a pipette with a tip of wide diameter. A repeating pipette may be used for these additions provided that the aliquots are added briskly to minimise the settling of cells.

Gently swirl the resulting cultures to mix the contents of the wells without cross–contaminating wells.

Remark: It is also valid to test the sample in a total culture volume of 200 µl. Take 40 µl of sample dilution, 110 µl medium and 50 ul of a cell solution (cell concentration 4 x 10^6 viable cells/ml)

Incubate the cultures without vibration (to allow the cells to settle) at 37°C for 16 – 24h in an atmosphere of 5% CO_2 in humidified air.

Harvesting and storage of supernatants.

At the end of the cell culture incubation the supernatant is harvested as follows:
Transfer from each well ca. 140 µl (or more if possible) of the supernatant above the MONOMAC 6 cells to the corresponding well a new 96-wells cell culture plate. Do not disturb the cells. The use of an eight channel pipette will facilitate the transfer. Change the pipette tips between each column transfer.
Cover the plate with the supernatants with a seal. Store the plate at –70 °C (or below) until required or proceed with the detection of IL6. When the plates are stored at –20 °C, determination of IL6 should be performed within two weeks.

7.4 Detection of IL–6 in the supernatant medium by ELISA

Immunoreactive IL–6 in aliquots of the cell culture fluid (cell–conditioned medium) is quantified using a validated ELISA.
Two validated IL-6 ELISA's are described in the ANNEX. (The NOVARTIS ELISA for IL–6 and the Human IL–6 ELISA kit (CLB, Amsterdam, The Netherlands, code M1916) Other validated ELISAs may be substituted.

8. DATA ANALYSIS

All experiments are run with four replicates of the test compound on one plate. A standard curve in quadruplicate is included, using the International Standard for Endotoxin (calibrated in IU), ranging *as appropriate* from 0.25 IU/ml up to 4 IU/ml *or 0.125 IU/ml up to 2 IU/ml*.
The endotoxin value of each replicate is calculated from the endotoxin calibration curve of the IS for endotoxin plotted against the final OD at 450 nm, applying the 4-parameter logistic model and expressed as endotoxin-equivalents/ml (EU/ml).

8.1 Assay acceptance criteria

The mean OD at 450 nm is calculated within a quadruplicate (×4). One datapoint per quadruplicate may be designated as an outlier, as determined using Dixon's test (P>0.05), and removed from the data set used for calculation.
The positive control (final concentration of the spike when diluted in saline) should be between ± 20% of the theoretical value (expected concentration).
The OD at 450 nm of the positive product control (final concentration of the spike diluted in product) should be over 50% of the positive control and below the 200% of the positive control (50–200% endotoxin spike recovery). The OD at 450 nm of the positive controle should be higher then the limit of quantification (LOQ).
The LOQ is defined as the mean OD at 450 nm of the R0 (Negative Cell Control) + 10xSD mean OD at 450 nm of the R0.
The OD at 450 nm of the negative control (blank) should be less then 0.200.
In the LPS standard curve the mean OD at 450 nm of the Rn should be less then the mean OD at 450 nm of the Rn+1. A minimum of 4 data points is needed for a valid reference curve.

8.2 Prediction model

For the purpose of the validation study, a sample is considered positive when the mean OD at 450 nm of the tested product exceeds the mean OD at 450 nm of that of the positive product control.

9. HEALTH SAFETY AND ENVIRONMENT

Human material

Human material should be treated as biologically hazardous and all work using human material is to be carried out according to laboratory safety procedures.

Cultures of human material should be treated as biologically hazardous waste and disposed of according to the laboratory safety procedures.

Bacterial endotoxin

Bacterial endotoxin is, as its name indicates, a toxic agent and should be handled with care.

Precautions: Cover open cuts before use. Do not get in eyes, on skin, on clothing. Avoid inhaling. Keep container closed.

First Aid: In case of contact, immediately flush eyes or skin with plenty of water for at least 15 minutes. If inhaled, remove to fresh air. If not breathing, give artificial respiration, preferably mouth–to–mouth. If breathing is difficult, give oxygen.

Effects of skin absorption can include fever, headache and hypotension.

Effects of inhalation can include fever, headache and hypotension.

Effects of ingestion – adverse effects are unlikely since ingested endotoxin is rapidly detoxified.

ELISA Substrate: TMB ***(Suspected mutagen, wear gloves when handling).*** Store and use in accordance with manufacturer's instructions.

10. REFERENCES

Bleeker, W.K., de Groot, E.M., den Boer, P.J., Biessels, P.T., Aarden, L.A. and Bakker, J.C. (1994) Measurement of interleukin–6 production by monocytes for in vitro safety testing of hemoglobin solutions. Artif Cells Blood Substit Immobil Biotechnol. 22, 835–40.

Dinarello, C.A., Gatti, S. and Bartfai, T. (1999) Fever: links with an ancient receptor. Curr Biol. 9, R147–50.

Dinarello, C.A., O'Connor, J.V., LoPreste, G. and Swift, R.L. (1984). Human leukocyte pyrogen test for detection of pyrogenic material in growth hormone produced by recombinant Escherichia coli. J. Clin. Microbiol. 20, 323–329.

Duff, G.W. and Atkins, E. (1982) The detection of endotoxin by *in vitro* production of endogenous pyrogen: comparison with amebocyte lysate gelation. J Immunol Methods 52, 323–332.

Fennrich, S., Fischer, M., Hartung, T., Lexa, P., Montag–Lessing, T., Sonntag, H.G., Weigandt, M. and Wendel. (1999) A Detection of endotoxins and other pyrogens using human whole blood. Dev Biol Stand. 101, 131–9.

Finney DJ: Statistical Method in Biological Assay, Third Edition. London, Charles Griffin and Company Ltd, 1978.

Gaines Das RE and Tydeman MS. (1982) Iterative weighted regression analysis of logit responses: A computer program for analysis of bioassays and immunoassays. Computer Programs in Biomedicine 15; 13–22.

Hansen, E.W. and Christensen, J.D. (1990) Comparison of cultured human mononuclear cells, Limulus amebocyte lysate and rabbits in the detection of pyrogens. J Clin Pharm Ther. 15, 425–33.

Levin, J. and Bang, F.B. (1964) A description of cellular coagulation in the Limulus. Bull. John Hopkins Hosp. 155, 337–345.

Mascoli, C.C. and Weary, M.E. (1979a) Limulus amebocyte lysate (LAL) test for detecting pyrogens in parenteral injectable products and medical devices: advantages to manufacturers and regulatory officials. J Parenter Drug Assoc. 33, 81–95.

Mascoli, C.C. and Weary, M.E. (1979b) Applications and advantages of the Limulus amebocyte lysate (LAL) pyrogen test for parenteral injectable products. Prog Clin Biol Res. 29, 387–402.

Poole, S., Selkirk, S., Rafferty, B., Meager, A., Thorpe, R. and Gearing, A. (1989) Assay of pyrogenic contamination in pharmaceuticals by cytokine release. Proceedings of the European Workshop on detection and quantification of pyrogen. Pharmeuropa, November 1989, 17–18.

Poole, S., Thorpe, R., Meager, A., Hubbard, A.R. and Gearing, A.J.H. (1988) Detection of pyrogen by cytokine release. Lancet 8577, 130.

Ray, A., Redhead, K., Selkirk, S. and Poole, S. (1990) Variability in LPS composition, antigenicity and reactogenicity of phase variants of Bordetella pertussis. FEMS Microbiology Letters 79, 211–218.

Taktak, Y.S., Selkirk, S., Bristow, A.F., Carpenter, A., Ball, C., Rafferty, B. and Poole, S. (1991) Assay of pyrogens by interleukin–6 release from monocytic cell lines. J Pharm Pharmacol 43, 578–582.

11. ANNEX

NOVARTIS IL–6 ELISA:

Materials, reagents.

Nunc-immuno MaxiSorp F96)
Mouse monoclonal anti–IL–6 antibody from clone 16 (Novartis)
Horseradish peroxidase conjugated sheep polyclonal anti–IL–6 antibody (Novartis)
3,3',5,5'–Tetramethyl benzidine (e.g. Fluka Cat. No. 87748)
Acetone (reagent grade)
Ethanol (reagent grade)
Phenol (e.g. Merck Cat. No. 100206)
Potassium hydroxide (reagent grade)
Sodium hydroxide (reagent grade)
Hydrochloric acid (reagent grade)
Sodium dihydrogen phosphate (e.g. Merck Cat. No. 106346)
Disodium hydrogen phosphate (e.g. Merck Cat. No. 106580)
Tris (hydroxymethyl) aminomethane (e.g. Fluka Cat. No. 93352)
Kathon MW/WT, Christ Chemie AG, Reinach, Switzerland
Albumin from bovine serum (e.g. Fluka Cat. No. 05480)
Citric acid monohydrate (e.g. Fluka Cat. No. 27490)
Sulphuric acid (2 M H_2SO_4, reagent grade)

Preparation of buffers

Coating Buffer
Dissolve 5.0 g sodium dihydrogen phosphate and 2.9 g disodium hydrogen phosphate in 400 ml distilled water. Use 1 N NaOH to adjust the pH to 7.5, and make up to 500 ml with distilled water.
Remains stable for 6 months at 2–8°C.

Blocking Buffer
Dissolve 12.1 g Tris (hydroxymethyl)aminomethane in 400 ml distilled water.
Add 0.1 ml Kathon MW/WT. Use 4 M HCl to adjust the pH to 7.5.
Add 5.0 g BSA. Add distilled water to make up to 500 ml.
Remains stable for 6 months at 2–8°C.

Stopping Solution
Add 26.6 ml H_2SO_4 to 500 ml distilled water.

Wash Solution
Add 1 ml Tween –20 to 2000 ml of demineralised water.

Dillution Buffer
Dissolve 2.1 g Tris(hydroxymethyl)aminomethane in 400ml distilled water.
Add 0.1 ml Kathon MW/WT, 0.5 g phenol and 25 ml heat–inactivated (30 minutes at +56°C) foetal bovine serum.
Mix to dissolve the substances, then adjust the pH to 7.5 with 4 M HCl. Make up to 500 ml with distilled water.
Remains stable for at least 6 months at 2–8 C.
In the absence of the stabilisers Kathon and phenol the stability is only 1 day.

TMB Solution #

Dissolve 240 mg TMB in 5 ml acetone. Add 45 ml ethanol and 0.3 ml Perhydrol (30 % H_2O_2).
Remains stable for at least 6 months at 15–25°C when sealed and protected from light.

Substrate Buffer[#]
Dissolve 6.3 g citric acid monohydrate in 800 ml distilled water.
Adjust the pH to 4.1 by adding 4 M KOH.
Make up to 1000 ml with distilled water and add 0.2 ml of Kathon MW/WT.

Remains stable for about 6 months at 15–25°C.
In the absence of the Kathon the stability is only 1 day.

([#] TMB solution and substrate buffer may be replaced by a TMB ready-to-use substrate system (e.g. Sigma, T8665)).

Coating of IL–6 ELISA plates

For the NOVARTIS IL–6 ELISA, dilute the coating anti–IL–6 antibody (Clone 16) with coating buffer to 2.5 µg/ml and swirl to mix, e.g. 1 mg of antibody in 400 ml of coating buffer. Add 200 µl to each well of a 96–well plate (Nunc–Immuno MaxiSorp F96). Stack the microtitre plates and allow to stand in the dark at 15–25°C for 16–24 h.

Aspirate and discard the coating solution. Wash the coated plate 3 times with demineralised water and tap out onto absorbent material, e.g. paper towel. Pipette 200 µl of blocking buffer into each well to block the residual protein–binding capacity of the coated plates. Seal the microtiter plates with adhesive film and store in a humidified atmosphere at 2–8°C (shelf life: two months).

DETECTION OF IL-6 IN THE CELL CONDITIONED MEDIUM BY ELISA

Equilibration of reagents

Bring assay reagents to room temperature before proceeding.
N.B. No IL-6 standard is used.

Addition of samples

Immediately prior to adding standards and samples, empty the blocking buffer from the ELISA plate(s) and expel any remaining fluid by tapping the inverted plate onto absorbent material, e.g. paper towel.
Make the supernatant more homogeneous by aspirating and expelling three times before transferring the test- amount of supernatant from each of the wells of columns 1 – 12 of the tissue culture plate into the corresponding wells on the ELISA plate – see Template 1 The use of an eight channel pipette will facilitate the mixing and transfer. Change the pipette tips between each column transfer.

Void wells are filled with Dilution Buffer.

Addition of (2^{nd}) antibody–HRP conjugate (POD)

Add 200 µl of detection antibody POD (horseradish peroxidase conjugated to sheep anti–IL–6 antibodies: stable for at least 6 months at 2–8°C) pre–diluted with dilution buffer (usually 1/200 to 1/500, as determined in optimisation experiments) to each well, seal the plates with adhesive film, and allow to stand for 2–3 hours at 20–25°C. (100 ml of diluted POD is sufficient for 4 ELISA plates.)

After incubation, wash plate three times with about 250 µl per well wash solution and then three times with demineralised water. Empty plate and expel any remaining fluid by tapping the inverted plate onto absorbent material, e.g. paper towel.

Addition of substrate solution and reading of optical densities

Prepare the substrate solution shortly before use. Transfer 90ml of substrate buffer to a plastic bottle, add 4.5 ml of TMB solution and mix.

Pipette 200µl of substrate solution into each well. After 10–15 minutes, stop the enzyme reaction by adding 50 µl/well of stopping solution. Wipe the back of the microtitre plates with a clean tissue, then measure the absorbance at 450 nm in an ELISA plate reader using a 540–590nm corrective filter. Subtract the values of the measurement with the corrective filter from values measured with the 450 nm filter.

IL-6 ELISA kit (CLB, Amsterdam, The Netherlands, code M1916)

The reagents provided in the ELISA kit used are:
- Coating antibody, 100-fold concentrated
- Blocking reagent, 50-fold concentrated
- Biotinylated IL-6 antibody, 100-fold concentrated
- Streptavidin-HRP conjugate, 10,000-fold concentrated
- Dilution buffer, 5-fold concentrated
- Microtiter plates + lid (Nunc-Immuno Maxisorp F96)
- Plate seals

Additional materials needed are:
Carbonatebuffer (pH 9.6)
Ethanol 96%
Tween 20
Phosphate buffered saline
Sulphuric acid (2 M H_2SO_4, reagent grade)
3,3, 5,5 Tetramethylbenzidine
Peroxide (30 % H_2O_2)
sulphuric acid (2M)
sodium-acetate (1.1 M, pH 5.5)
Instead of the last four items, a TMB ready-to-use substrate system can be applied (e.g. Sigma, T-8665)

Preparation of reagent solutions used in the IL-6 ELISA
The following solutions are prepared prior to each test:

Washing buffer
Add 50 µl Tween 20 (or 500 µl 10% Tween 20) to 1000 ml PBS.

Dilution buffer
The ELISA kit contains one bottle with 5-fold concentrated dilution buffer. Calculate the quantity of dilution buffer required and prepare a working-strength dilution by diluting the concentrated buffer 1 in 5 in distilled water.

Substrate
Dissolve 6 mg TMB in 1 ml 96% ethanol. Add 1.2 ml sodium-acetate (1.1 M, pH 5.5) and 0.2 ml TMB solution to 10.8 ml of distilled water. Add 2.4 µl H_2O_2 directly prior to use.
(alternatively: an equivalent TMB substrate system can be applied)

Performing the ELISA
Centrifuge all vials (except the blocking reagent) before use (1 min. at 3000 g).
Preferably an ELISA-plate shaker is used during the incubations (except during the coating of the plate))

Coating of the plate
Dilute 120 µl of the coating antibody with 12 ml of carbonate buffer (= 1/100). Add 100 µl of this dilution to each well of the microtiterplate and incubate overnight at RT.

Wash the plate four times using washing buffer (platewasher).

Blocking of the plate

Dilute 500 µl blocking reagent with 25 ml PBS (= 1/50). Add 200 µl of this dilution to each well of the microtiterplate and incubate during 1 hour at RT. In the meantime bring the storage plate with the supernatant to RT.
Wash the plate four times using washing buffer.

Addition of the samples (supernatants)
See: -template 1 above

Homogenise the supernatant in the storage plate before transferring the test-amount of supernatant from the storage plate to the ELISA. Add 80 µl dilution buffer and 20 µl of the sample to the assigned wells of the microtiterplate. As a controle for the performance of the ELISA, one of the sample can be replaced with dilution buffer. Incubate for 1 hour at RT (Plate shaker).
Void wells are filled with Dilution Buffer.
N.B. The pate design does not allow the use of the IL-6 standard.

Wash the plate four times using washing buffer.

Conjugate 1: Biotinylated IL-6 antibody
Add 120 µl biotinylated IL-6 antibody to 12 ml dilution buffer (= 1/100).
Add 100 µl of the conjugate dilution to each well and incubate for 1 hour at RT. (Plate shaker)

Wash the plate four times using washing buffer.

Conjugate 2: Streptavidine-Peroxidase
Add 3 µl streptavidine-HRP conjugate to 30 ml dilution buffer (= 1/10,000).
Add 100 µl of the conjugate dilution to each well and incubate for half an hour at RT. (Plate shaker)

Wash the plate four times using washing buffer.

Substrate solution
Add 100 µl of substrate solution to each well and incubate 15 minutes at RT

Stop solution
Add 100 µl of sulphuric acid (2M) to each well. Measure the absorbance at 450 nm with an ELISA plate reader.

[This Page Intentionally Left Blank]

Appendix B

ECVAM Response to ICCVAM Questions

ECVAM Response to ICCVAM Questions B-1

[This Page Intentionally Left Blank]

EUROPEAN COMMISSION
DIRECTORATE GENERAL JRC
JOINT RESEARCH CENTRE
Institute for Health and Consumer Protection JRC
European Centre for the Validation of Alternative Methods (ECVAM)

Ispra, 24 Mar. 06

Response to the ICCVAM-PWG Questions for ECVAM regarding the five in vitro pyrogenicity test method submissions

Background

The concept of pyrogen testing employing the human fever reaction has been first proposed 25 years ago by Dinarello et al., when the mediators of the fever reaction were identified and could be used as a measure of the response to pyrogens. Since, a variety of test systems employing this reaction has been proposed and used in more than 100 institutions. Over the last decade, about 5 million $/€ of public funding by the EU, the German BMBF, ZEBET and ECVAM enabled the standardization and formal validation of the most eminent approaches.
Following an ECVAM workshop in 2000 (Hartung et al. 2001), a validation study was organized, which targeted solely the replacement of the rabbit animal test. The lower limit of detection of these assays compared to the rabbit and the fact that the novel assays can be used in a quantitative manner was not addressed in this validation study. The study also addressed specifically products, which are tested in rabbits, i.e. intravenous parenterals with a focus on those not testable in the Limulus amoebocyte lysate (LAL) assay/bacterial endotoxin test (BET). Since both the rabbit test and LAL/BET are mainly reactive to Gram-negative endotoxin, the study focused on this predominant endotoxin. However, outside the validation study a wealth of scientific information, especially on the human whole blood pyrogen tests, is available, which was referred to in the dossiers.
The developing laboratories had been held responsible for the content of the individual dossiers, which were commissioned by ECVAM. They were contacted to provide information to respond to the questions of the ICCVAM-PWG. The respective responses are compiled below. Editorial changes as suggested will be carried out on a short term.

1. In the BRDs, the accuracy evaluation is based on the application of a prediction model derived from historical in vivo rabbit data. However, if the proposed test methods are intended as replacements for the current pyrogenicity test methods (i.e., rabbit pyrogen test; BET), should not the accuracy analysis be a direct comparison of the proposed in vitro methods with both of these reference test methods? Can you provide data from such a comparison for review?

Due to ethical and legal reasons, it was not possible to perform the rabbit pyrogen test in parallel to the in vitro methods. The development of the prediction model was therefore based on historical rabbit data. Since rabbit strains differ in their sensitivity, data generated in the most sensitive rabbit strain were used. 60 years of pharmcopoeial use of the rabbit assay has shown, that the limit concentration is effectively protecting humans. The availability of a WHO reference material also allowed basing the study on a historic comparison.

On the basis of the determined rabbit fever threshold (Hoffmann S, Luderitz-Puchel U, Montag-Lessing U and Hartung T. Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modelling, J. Endotoxin Res. 2005, 11:25-31), it was possible to model the performance of these rabbits, when testing the samples of the validation study assuming no additional interference of the samples. The sensitivity of the rabbit pyrogen test is 57.9% and the theoretical specificity of the rabbit pyrogen test is 88.3%. Unfortunately these numbers have not been correctly reported in the BRD, in sections 4.2 and 4.6., where the last sentence reads " .. sensitivity of the rabbit pyrogen test is 75.04% and the theoretical specificity of the rabbit pyrogen test is 95.80%.

Joint Research Centre · I-21020 Ispra (VA), Italy · TP 202
Telephone: direct line (+39-0332) 785939, · Telefax: 786297
E-mail Internet:thomas.hartung@cec.eu.int
WWW: http://www.jrc.cec.eu.int/

Comparisons between rabbit and human tests were also carried out by establishing a rabbit whole blood assay (Schindler S et al. Comparison of the reactivity of human and rabbit blood towards pyrogenic stimuli. ALTEX 2002, 20:59-63). The study reflected the slightly lower sensitivity of rabbit towards endotoxins and discrepancies for non-endotoxin pyrogens.

Based on historic rabbit results, it has been shown that the enormous differences in potency of endotoxins (4 log orders) from different bacterial species are reflected in both the rabbit and the WB/IL-1 but not in the BET (Fennrich S et al. Detection of endotoxins and other pyrogens using human whole blood. Dev. Biol. Standards, 1999, 101:131-139). The BET can thus not be used as a quantitative measure of endotoxicity to mammals if the bacterial species is not known. This has been recently expanded (Dehus O, Hartung T, and Hermann C. Pseudomonas endotoxin, a prominent contamination of water samples, is strongly overestimated in the Limulus Amebocyte Lysate Assay with regard to its pyrogenicity. J Endotox Res 2006, in press.) to endotoxins of Pseudomonas, which are overestimated by the BET by 2-3 log orders.

Although, the BET was not addressed in the validation study, it has several well-known limitations, which might in the future lead to additional uses of the novel tests:

- BET is restricted to Gram-negative endotoxin and misses those from Gram-positive bacteria or fungi
- BET it is disturbed by many components such as endotoxin-binding proteins (ample e.g. in blood-derived medicines), lipophilic substances, glucans present in herbal medicinal products or derived from cellulose filters; in fact, about 90% of LAL testing is done on water samples only due to these limitations
- BET does not reflect the biological potency of different endotoxins in humans (discrepancies up to 10.000fold)
- BET consumes animals (several recent studies indicate 15% mortality of bled animals to be seen in the light of about 60% reduction in horseshoe crab landings between 1998 and 2003)
- BET does not work for solid materials (medical devices) or cellular therapies without difficult extraction procedures
- BET does not work for air-born pyrogens increasingly recognized as a health threat, because of the restriction to Gram-negative endotoxins, not reflected endotoxin potency and impossibility to measure directly on air filters

Within the validation study, two samples of a human serum albumin (HSA) were provided to the developing labs. One of these was essentially pyrogen-free, the other one was a contaminated lot that was associated with adverse reactions in recipients. The labs themselves were responsible for the testing and supply of the results. Two labs, responsible for the THP-1/TNF and the PBMC/IL-6 methods, did not report results for this phase. Generally, the clean lot was used to determine an interference free dilution of the HSA. Then, although dilution response curves were also produced for the contaminated lot, the response of this dilution was assessed by means of an LPS standard curve. The same approach was taken with the two lots of Gentamicin that were provided. Again, one of these was clean and one was associated with adverse reactions in recipients. The results for the HSA and Gentamicin with regard to the discrimination of the pyrogen-free and contaminated lots are presented in the tables.

HSA	THP-1-Neo	MM6-IL6	blood-IL6	blood-IL1	PBMC-IL6
Interference free dilution	1:4	1:2	1:4 – 1:8	1:4	1:10
Discrimination of lots	+	+	+	-	+

Table: Results of HSA-testing for four methods

For the HSA the THP-1-Neo, MM6-IL6, blood-IL6 and PBMC-IL6 were able to distinguish the contaminated lot from the pyrogen-free lot.

Gentamicin	THP-1-Neo	MM6-IL6	blood-IL6	blood-IL1	PBMC-IL6
Interference free dilution	not testable	1:4	1:1 (2 donors)*	1:10	1:16
Discrimination of lots	-	+	+	+	+

Table: Results of Gentamicin-testing for four methods (* one donor not suitable)
For Gentamicin only the THP-1-Neo could not distinguish the lots because of severe interference. Although the interference free dilutions differed, the other four methods could discriminate the pyrogen-free and the contaminated lot.

Outside the validation study, direct comparisons have been reported for the blood-IL-1 test (referred to as IVPT) by the German National Control Authority, the Paul-Ehrlich-Institute (PEI), to the European Pharmacopoeia:

In vitro Pyrogen Test (IVPT) in comparison to Rabbit Pyrogen Test (RPT) and Limulus Test (LAL)

a) Fever reactions caused by a batch of Human Serum Albumin
(negative in RPT as well as in LAL test)

After application of a defined batch of Human Serum Albumin, fever reactions have been reported to the PEI. The batch has been withdrawn from the market. The manufacturer had tested the product in RPT as a release criterion with negative result.

The PEI examined samples of the batch in RPT, in IVPT (Whole Blood Pyrogen Test), and in LAL. Negative Albumin batches of the same manufacturer served as controls.

RPT as well as LAL remained negative. The results of Whole Blood Pyrogen Test (5 different donors, at least 15 repetitions per donor) are shown in table2.

Table: Incriminated Human Serum Albumin

Donor	Incriminated batch IL-1 (pg/ml)	Control batch IL-1 (pg/ml)	Quotient incriminated/control
1	79.0	4.0	19.75
2	14.1	3.9	3.61
3	44.3	15.0	2.95
4	20.9	14.9	1.4
5	71.9	3.9	18.44
Mean	46.04	8.34	5.52

Remarks
A) In routine IVPT of PEI, pools from blood of 5 donors are used equalising the differences in donor reactions. The incriminated batch, negative in LAL as well as in RPT, had been clearly detected in IVPT.
B) There is an individual susceptibility towards Non-endotoxin pyrogens in humans in contrast to endotoxin that leads to fever reactions in almost all recipients.

b) Coagulation Factor VIII Concentrates
(negative in RPT and positive in IVPT)

In an early pilot study, five different Factor VIII Concentrates (one batch each) have been tested in RPT and in IVPT (Whole Blood Pyrogen Test) in parallel. The results are shown in table 2.

Table: Pilot study Factor VIII

Product	Whole Blood Pyrogen Test			Rabbit Pyrogen Test	
	IL-1 (pg/ml)	cut off (pg/ml)	result	temperature sum of 3 rabbits	result
A	130.4	18.6	positive	0.71 °C	negative
B	159.3	27.8	positive	0.70 °C	negative
C	32.6	32.6	negative	0.45 °C	negative
D	33.5	28.8	positive	1.6 °C*	negative
E	129.7	15.5	positive	0.37 °C	negative

* 6 rabbits (repetition)

It has to be taken into account that in case of F VIII – because of the small injection volume of 50 IU F VIII per kg corresponding to 500 µl per kg – the Rabbit Pyrogen Test is less sensitive (20 IU LPS/ml) than the Whole Blood Pyrogen Test (detection limit 0.25 to 0.5 IU LPS/ml).

It should be pointed out that there was a discussion in group 6B concerning the increase of injection volume of F VIII preparations in RPT even for protection of bleeding patients requiring greater amounts of F VIII.

c) Pyrogenic batch of a plasma derivative
(positive in RPT, negative in LAL test)

One batch of a partially purified, immunoglobulin containing plasma derivative was tested positive in RPT and negative in LAL by PEI. By the time the IVPT had been implemented, there was no further sample of the product available.

d) Factor VIII Concentrates, comparison study with spiked products.

Factor VIII concentrates from 6 different manufacturers (Immunate STIMplus 1000, Haemoctin SDH 1000, Octonate 1000, Fandhi 1000 IE, Beriate P 1000, and Haemate HS 1000) were tested (3 lots each) in a comparative study between RPT and IVPT (according to EP). The 18 preparations were spiked with two different concentrations of endotoxin (WHO Standard) and, thereafter, tested on the same day in parallel. Un-spiked preparations served as controls. All together, 162 rabbits were involved into the study.

The results are shown in table 3. Applying 5 IU endotoxin per kg body weight of the rabbits (representing the fever threshold of most sensitive rabbit strains) almost all rabbits responded with a temperature increase in the range allowing a repetition of the test. No positive result could be seen whereas the IVPT produced positive results in every experiment. After application of the 3-fold dose LPS in Factor VIII, the RPT identified 11 of 18 preparations clearly positive, in 7 cases the repetition range was achieved. Again, all samples in IVPT were positive. A false negative result was observed in one experiment of IVPT (1 of 4 donors); the repetition of the experiment remained negative.

Table: Comparison study Factor VIII (18 batches, 162 rabbits)

Endtotoxin (WHO Standard)		Rabbit Pyrogen Test 3 animals per test			Whole Blood Pyrogen Test n = 4 donors each	
Rabbit	IVPT	positive	repetition possible	negative	positive	negative
IU LPS / kg (1 ml / kg)	IU LPS / ml	> 2.65 °C	> 1,15 < 2.65 °C	< 1.15 °C		
0	0	0	0	18	0 (1)*	72 (71)*
5	5	0	17	1	72	0
15	15	11	7	0	72	0

* testing the same sample, the blood of 3 donors remained negative, the blood of 1 donor reacted slightly positive in the first experiment, the repetition was negative

e) Human Serum Albumins, comparison study with spiked products

Human Serum Albumins of various protein concentrations from five different manufacturers (5 % Immuno, 20 % Immuno, 20 % Kabi, 5 % Biotest, 20 % Biotest, 25 % Biotest, 5 % Centeon, 20 % Centeon, 5 % DRK BaWue, 20 % DRK BaWue) were tested (3 batches each, in case of 25 % Biotest only 2 batches were available) in a comparative study between RPT and IVPT (according to EP). The 29 preparations were spiked with two different concentrations of endotoxin (WHO Standard) and, thereafter, tested on the same day in parallel. Un-spiked preparations served as controls. All together, 261 rabbits were involved into the study.

The results are shown in table 4 (see page 4). Only 5 batches containing the borderline endotoxin concentration could be defined as positive by the RPT whereas the IVPT identified all samples. Applying the double dose, in 21 of 29 cases a clear positive result could be observed in the RPT, in 8 cases the repetition range was achieved.

Table: Comparison study Human Serum Albumin (29 batches, 261 rabbits)

Endtotoxin (WHO Standard)		Rabbit Pyrogen Test 3 animals per test			Whole Blood Pyrogen Test	
Rabbit	IVPT	positive	repetition possible	negative	positive	negative
IU LPS / kg	IU LPS / ml	> 2.65 °C	> 1,15 < 2.65 °C	< 1.15 °C		
0	0	0	0	29	0	29
5	5	5	23	1	29	0
10	10	21	8	0	29	0

It has to be pointed out that in most of the RPTs the maximal permitted injection volume of 10 ml per kg body weight has been used, which represents “worst case sensitivity conditions” of RPT (except for HSAs of 20 % and 25 % protein content where the EP allows lower volumes only). Thus, the sensitivity limit of RPT has been met in this study.

f) Fever reactions caused by an infusion solution

An infusion solution containing gelatine (release criterion LAL) induced adverse fever reactions in hospitals. The manufacturer withdrew the incriminated batches from the market and reinvestigated

them for endotoxin and, additionally, for pyrogenicity in RPT. The company observed LAL negative but RPT positive results in one batch. However, the most interesting batch which caused fever in patients remained negative in LAL as well as in rabbits. They were blinded by the manufacturer, sent to PEI, and analysed in Whole Blood Pyrogen Test. Furthermore, the PEI asked for blinded non-incriminated control batches.

Table 5: Incriminated infusion solution containing gelatine

batch	LAL test	rabbit test	fever in patients	Whole Blood Pyrogen Test		
				IL-1 (pg/ml)	IL-6 (pg/ml)	TNF □ (pg/ml)
A	negative	negative	no	8.5	28.0	28.2
B	negative	positive	yes	142.6	654.4	67.6
C	negative	negative	yes	421.5	9444.0	116.7
cut off:				32.6	127.6	43.6

The results are summarised in table 5. The incriminated batches could be identified very clearly in IVPT even the sample remaining negative in RPT. The fever causing substances have to be seen as non-endotoxin pyrogens not detectable in LAL. At least one of them is not pyrogenic for rabbits but for humans. The IVPT is basing on a “homogeneous indicator system”, the activation of human monocytes as central mediators of fever reaction. In conclusion, the IVPT indicates in vivo reactions of humans whereas the rabbit is not able to image every non-endotoxin pyrogen.

The results are in concordance with adverse fever reactions by an antibiotic (gentamycin) happened on the US market some years ago which had negative LAL test as release criterion also. The drug produced unclear results in RPT but could be tested as positive in IVPT (personal communication from Stephen Poole, NIBSC).

Remark: Unfortunately, the manufacturer of the infusion solution does not agree in publication of the results.

2. The proposed applicability domain of the test methods covers both endotoxin and non-endotoxin pyrogens. How do the validation studies included in the BRDs support both indications when there is no data provided to indicate their utility for non-endotoxin pyrogens? Can you provide data for any of the proposed test methods tested with non-endotoxin pyrogens?

Within the validation study, eleven substances provided to the developing labs of the new methods were tested. For each substance the smallest concentration (in ng/ml) inducing a response in the respective method was reported. In the table a summary of the results is presented.

	WBT/IL1	WB/IL-6	PBMC/IL-6	MM6/IL-6	THP-1/Neo
Curdlan	1000	1000	100	1000	1000
Glucan-Barley	Negative	negative	negative	negative	negative
Glucan-Yeast	not done	negative	negative	negative	negative
Zymosan	Negative	10000	10000	negative	10000
PHA-L	100	10000	100	100	1000

PHA-E	Negative	negative	negative	10000	negative
Lipid A	10000	1000	10000	1000	negative
Glucan STD	Negative	negative	negative	negative	undiluted
Endotoxin-C	4	40	0.4	4	4
Endotoxin-G	0.4	40	4	4	4
LTA	5000	500	5000	6250	1000

Table: Smallest concentration [ng/ml] or dilution of substances active in the respective method

The three glucans, which give false positive results in the BET, were not active at the highest concentration of 10000 ng/ml assessed in the five new methods in which they were tested. The Curdlan, which reacted in the BET, induced a cytokine response in the new methods generally only at a concentration of some 1000 ng/ml. Also, Zymosan and monophosphoryl-Lipid A were inactive or active only in very large doses in the new methods. The results for the PHA-L differed slightly between the methods. It was most reactive in the blood-IL1, the PBMC-IL6 and the MM6-IL6. However, the PHA-L was more reactive than the PHA-E, which was largely inactive. The potencies of the Pseudomonas endotoxins C and G differed somewhat in the different methods. The LTA was most active in the blood-IL6 and the THP-1-Neo methods. Taken together, the assays did not react to the LAL false-positive glucans and curdlan, but reacted to the LAL-false negative LTA.

WB/IL-1

The developing laboratory of the WBT tests has in fact been using this assay to identify and purify the Gram-positive counterpart, i.e. lipoteichoic acid (LTA), to Gram-negative lipopolysaccharide (LPS) in a series of studies:

1. Figuero-Perez I, Stadelmaier A, Morath S, Hartung T and Schmidt RR. Synthesis of structural variants of Staphylococcus aureus lipoteichoic acid (LTA). Tetrahedron 2005, 16:493-506.
2. Grandel U, Hopf M, Buerke M, Hattar K, Heep M, Fink L, Bohle RM, Morath S, Hartung T, Pullamsetti S, Schermuly RT, Seeger W, Grimminger F and Sibelius U. Mechanisms of cardiac depression caused by lipoteichoic acids from Staphylococcus aureus in isolated rat hearts. Circulation 2005, 112:691-698.
3. Grangette C, Nutten S, Palumbo E, Morath S, Hermann C, Dewulf J, Pot B, Hartung T, Hols P and Mercenier A. Enhanced anti-inflammatory capacity of a Lactobacillus plantarum mutant synthesizing modifed teichoic acids. Proc. Natl. Acad. Sci. U S A 2005, 102:10321-10326.
4. Henneke P, Morath S, Uematsu S, Weichert S, Pfitzenmaier M,Takeuchi O, Mueller A, Poyart C, Akira S, Berner R, Teti G, Geyer A, Hartung T, Trieu-Cuot P, Kasper DL and Golenbrock DT. Role of lipoteichoic acid in the phagocyte response to group B Streptococcus. J. Immunol. 2005, 174:6449-6455.
5. Hermann C, von Aulock S, Dehus O, Keller M, Okigami H, Gantner F, Wendel A, and Hartung T. Endogenous cortisol determines the circadian rhythm of LPS- but not LTA-inducible cytokine release. Eur J Immunol 2005, in press.

6. Hoebe K, Georgel P, Rutschmann S, Du X, Mudd S, Crozat K, Sovath S, Shamel L, Hartung T, Zahringer U and Beutler B. CD36 is a sensor of diacylglycerols. Nature 2005, 433:523-527.

7. Kinsner A, Pilotto V, Deininger S, Brown GC, Coecke S, Hartung T and Bal-Price A. Inflammatory neurodegeneration induced by lipoteichoic acid from Staphylococcus aureus is mediated by glia activation: the role of nitrosative and oxidative stress, mitochondrial dysfunction and caspase activation. J Neurochem 2005, 95:1132-1143.

8. Morath S, von Aulock A and Hartung T. Structure/ function relationships of lipoteichoic acids. J. Endotoxin Res. 2005, in press.

9. Dahle MK, Overland G, Myhre AE, Stuestol JF, Hartung T, Krohn CD, Mathiesen O, Wang JE and Aasen AO. The phosphatidylinositol 3-kinase/protein kinase B signaling pathway is activated by lipoteichoic acid and plays a role in Kupffer cell production of interleukin-6 (IL-6) and IL-10. Infect Immun 2004, 72:5704-11.

10. Lotz S, Aga E, Wilde I, van Zandbergen G, Hartung T, Solbach W and Laskay T. Highly purified lipoteichoic acid activates neutrophil granulocytes and delays their spontaneous apoptosis via CD14 and TLR2. J. Leukoc. Biol. 2004, 75:467-477.

11. Lynch NJ, Roscher S, Hartung T, Morath S, Matsusshita M, Maennel DN, Kuraya M, Fujita T and Schwaeble WJ. L-ficolin specifically binds to lipoteichoic acid, a cell wall constituent of Gram-positive bacteria, and activates the lectin pathway of complement. J. Immunol. 2004, 172:1198-1202. (IF 6.7)

12. Mattsson E, Hartung T, Morath S and Egesten A. Highly purified lipoteichoic acid from Staphylococcus aureus induces procoagulant activity and tissue factor expression in human monocytes but is a weak inducer in whole blood-comparison with peptidoglycan. Infect. Immun. 2004, 72:4322-4326.

13. Triantafilou M, Morath S, Mackie A, Hartung T and Triantafilou K. Lateral diffusion of Toll-like receptors reveals that they are transiently confined within lipid rafts on the plasma membrane. J. Cell Sci. 2004, 117:4007-4014.

14. Triantafilou M, Manukyan M, Mackie A, Morath S, Hartung T, Heine H and Triantafilou K. Lipoteichoic acid and Toll-like receptor 2 internalization and targeting to the Golgi is lipid raft dependent. J. Biol. Chem. 2004, 279:40882-40889.

15. Zeidler D, Zahringer U, Gerber I, Dubery I, Hartung T, Bors W, Hutzler P and Durner J. Innate immunity in Arabidopsis thaliana: lipopolysaccharides activate nitric oxide synthase (NOS) and induce defense genes. Proc. Natl. Acad. Sci. U S A 2004, 101:15811-6.

16. Deininger S, Stadelmaier A, von Aulock S, Morath S, Schmidt RR, Hartung T. Definition of structural prerequisites for lipoteichoic acid inducible cytokine induction by synthetic derivatives. J. Immunol. 2003, 170:4134-4138.

17. Overland G, Morath S, Yndestad A, Hartung T, Thiemermann C, Foster SJ, Smedsrod B, Mathiesen O, Aukrust P, Aasen AO and Wang JE. Lipoteichoic acid is a potent inducer of cytokine production in rat and human Kupffer cells in vitro. Surg. Infect. 2003, 4:181-191.

18. Schröder NWJ, Morath S, Alexander C, Hamann L, Hartung T, Zähringer U, Göbel UB, Weber JR, Schumann RR. Lipoteichoic acid (LTA) of S. pneumoniae and S. aureus activates immune cells via toll-like receptor (TLR)-2, LPS binding protein (LBP) and CD14 while TLR-4 and MD-2 are not involved. J. Biol. Chem. 2003, 278:15587-94.

19. Schröder NWJ, Hermann C, Hamann L, Göbel UB, Hartung T, Schumann RR. High frequency of polymorphism Arg753Gln of the Toll-like receptor-2 (TLR-2) gene detected by a novel allele specific PCR. J. Mol. Med. 2003, 81:368-72.

20. Stadelmaier A, Morath S, Hartung T, Schmidt RR. Synthesis of the first fully active

lipoteichoic acid. Angew. Chemie Int. Ed. 2003, 42:916-920. auch: Stadelmaier A, Morath S, Hartung T, Schmidt RR. Synthese der ersten aktiven Lipoteichonsäure. Angew. Chemie 2003, 115:945-949.

21. Dalpke AH, Frey M, Morath S, Hartung T, Heeg K. Interaction of lipoteichoic acid and CpG-DNA during activation of innate immune cells. Immunobiol. 2002, 206:392-407.
22. Hermann C, Spreitzer I, Schröder NWJ, Morath S, Lehner MD, Fischer W, Schütt C, Schumann RR and Hartung T, Cytokine induction by purified lipoteichoic acids from various bacterial species – role of LBP, sCD14, CD14 and failure to induce interleukin-12 and subsequent interferon-γ release, Eur. J. Immunol., 2002, 32:541-551.
23. Jacinto R, Hartung T, McCall C and Li L. Lipopolysaccharide and lipoteichoic acid-induced tolerance and cross-tolerance: distinct alterations in IL-1 receptor-associated kinase. J. Immunol. 2002, 168:6136-6146.
24. Morath S, Geyer A, Spreitzer I, Hermann C and Hartung T. Structural decomposition and heterogeneity of commercial lipoteichoic acid preparation. Infect. Immun. 2002, 70:938-944.
25. Morath S, Stadelmaier A, Geyer A, Schmidt RR and Hartung T. Synthetic lipoteichoic acid from Staphylococcus aureus is a potent stimulus of cytokine release. J. Exp. Med., 2002, 195:1635-1640.
26. Yipp BG, Andonegui G, Howlett CJ, Robbins SM, Hartung T Ho M and Kubes P. Profound differences in leukocyte-endothelial cell responses to lipopolysaccharide versus lipoteichoic acid. J. Immunol. 2002, 168:4650-4658.
27. Lehner MD, Morath S, Michelsen KS, Schumann RR and Hartung T. Induction of cross-tolerance by LPS and highly purified lipoteichoic acid via different Toll like receptors independent of paracrine mediators. J. Immunol. 2001, 165:5161-5167.
28. Morath S, Geyer A, Hartung T. Structure/function-relationship of cytokine induction by lipoteichoic acid from Staphylococcus aureus. J. Exp. Med., 2001, 193:393-397.
29. Michelsen KS, Aicher A, Mohaupt M, Hartung T, Dimmeler S, Kirschning CJ and Schumann RR. The role of Toll-like receptors (TLRs) in bacteria-induced maturation of murine dendritic cells -PGN and LTA are inducers of DC maturation and require TLR2. J. Biol. Chem. 2001, 276: 22041-22047.
30. Opitz B, Schröder NWJ, Spreitzer I, Michelsen KS, Kirschning CJ, Hallatschek W, Zähringer U, Hartung T, Göbel UB and Schumann RR. Toll-like receptor (TLR)-2 mediates treponema glycolipid and lipoteichoic acid (LTA)-induced NF-kappa B translocation. J. Biol. Chem. 2001, 276:22041-22047.
31. van de Wetering JK, van Eijk M, van Golde LMG, Hartung T, van Strijp JAG and Batenburg JJ, Characteristics of surfactant protein A and D binding to lipoteichoic acid and peptidoglycan, two major cell wall components of Gram-positive bacteria. J. Inf. Dis. 2001, 184:1143-1151.

LTA and LPS were compared in the WB/IL-1:

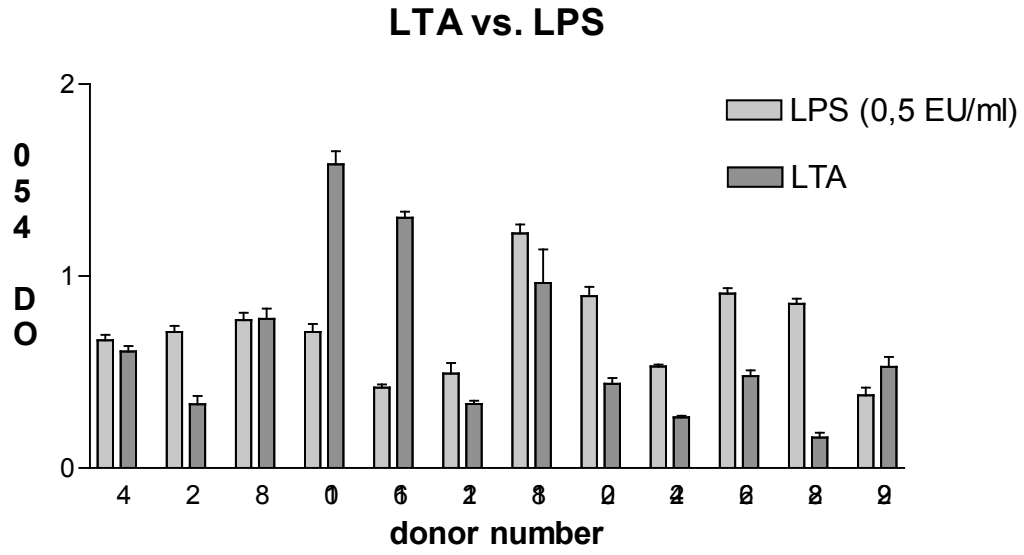


Fig. Reactivity of 12 donors towards LPS E. coli O113: H10 (0.5 EU/ml) and LTA from B. subtilis
The reactivity of 12 donors towards a challenge of 0.5 EU/ml LPS and the LTA was compared. In all cases, the stimuli tested clearly positive, although the donors did not necessarily react homogenously.

The studies were more recently expanded to fungi (Kindinger at al., A new method to measure air-borne pyrogens based on human whole blood cytokine response. J. Immunol. Meth. 2005, 298:143-153).
Fungi pose a particular health problem, since they or their spores are potentially neurotoxic, hepatotoxic, nephrotoxic and even teratogenic. The identification of pathogenic fungi in parenterals and, even more importantly, air, is therefore a particular challenge. Fungi in general proved to be highly active in the whole blood test, though differences between different strains exist (see figure). That the pyrogenic potency is not due to LPS is demonstrated in the subsequent figures employing the LPS inhibitor Polymyxin B.

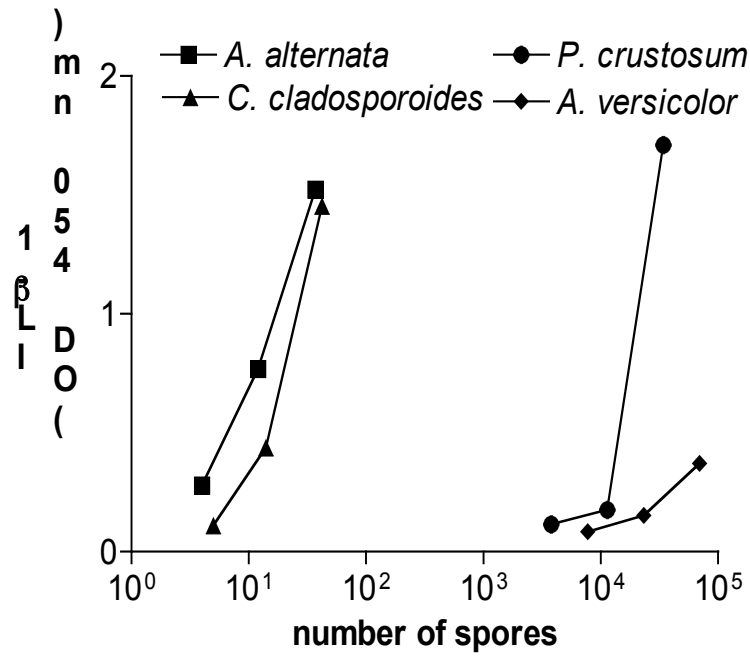
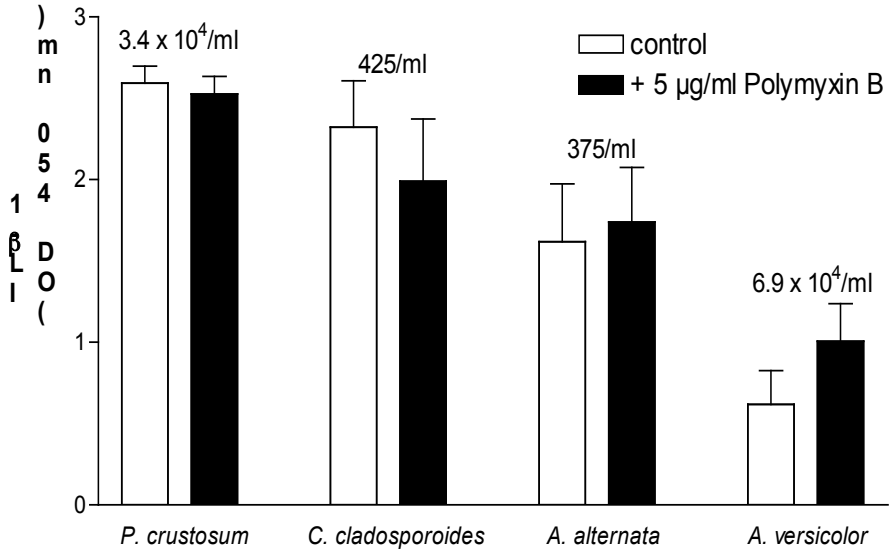


Figure: Fungal spores (Alternaria alternata, Cladosporium cladosporoides, Penicilium crustosum, Aspergillus versicolor) induce the release of IL-1□ by human whole blood.



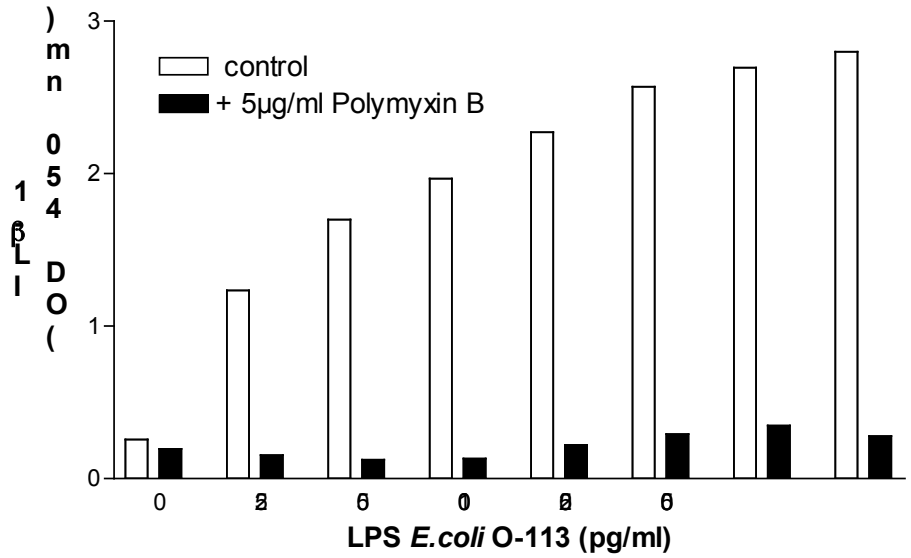


Figure 8: IL-1□ release in human whole blood in response to fungal spores (A) is not inhibited by 5 µg/ml polymyxin B, mean of 4 donors (±SEM), numbers above the bars indicate the spore counts employed. In contrast (B), the response to LPS is inhibited over a wide concentration range, mean of double values.

WB/IL-6
Carlin & Viitanen (In vitro pyrogenicity of diphtheria, tetanus and acelluar pertussis components of a trivalent vaccine, Vaccine 23, 3709-3715, 2005) report that IL6 release was triggered by toxoid deriving from Gram-positive bacteria *Corynebacterium diphtheriae* and highly purified lipoteichoic acid from Gram-positive bacteria *Bacillus subtilis* and *Staphylococcus aureus*.

MM6/IL-6
Preliminary tests performed with other substances than described in the BRD have shown that the MM6/IL-6 assay is also reactive with non-endotoxin pyrogens. However, these data are not easily evaluated as there are no standards for pyrogens (other than endotoxin) available.

3. The test substances included in the BRDs are all parenteral pharmaceuticals. However, if the proposed test methods are intended as replacements for the current pyrogenicity test methods (i.e., rabbit pyrogen test; BET) shouldn't the validation studies also include other relevant test materials (e.g., medical devices, biologics, etc.)? Can you provide data for any of the proposed test methods tested with other relevant materials? If such data are not currently available, do you intend to generate data in a subsequent phase of testing?

As explained in the background information, by purpose the study has targeted areas of use of the rabbit test not the BET. Notably, the validation study did include human serum albumin and factor VIII preparations. Furthermore, supportive information is available for some of the tests (see also question 1). Since it is common practice to validate pyrogen tests for every given product and the restrictions in resources, the validation study itself could not cover additional substances.

WB/IL-1
In collaboration with the Brazilian National Control Authority, snake venom sera have been studied.

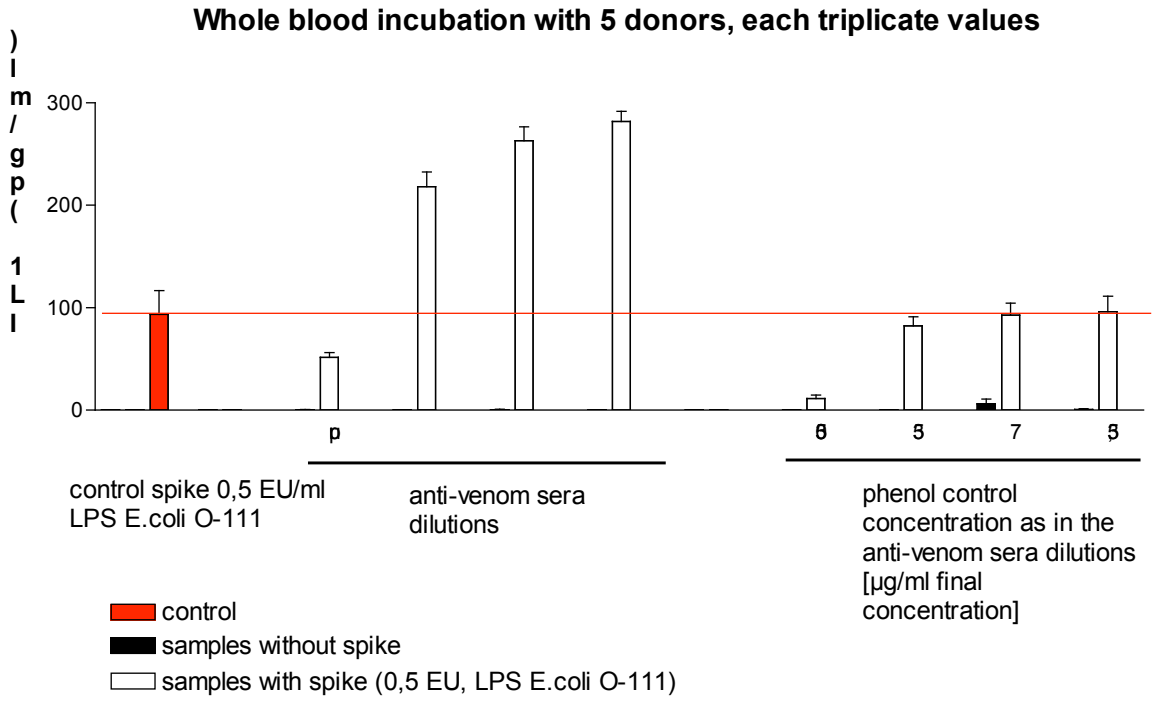


Figure: Spiked snake venom serum shows positive signal, though increased when compared to the control spike in saline.

Adaptation of the WB/IL-1 to biocompatible materials

Medical devices pose a particular problem since they cannot be examined directly with the abovementioned classical test systems. In order to be tested in the BET, an eluate has to be prepared, which is subsequently tested. This bears the risk to miss non-eluable, e.g. very hydrophobic, contaminations. Furthermore, the BET cannot detect non-endotoxin pyrogens. With the WBT, the material in question can be incubated directly and without any former treatment, provided, an adequate pyrogen-free control is provided. Apart from that their diversity with regard to size, form, material and form of application challenges biological assays demanding individual approaches. Products with direct (blood bags, needles) and indirect (swabs, gloves) contact to the blood circulation can have serious impact on the organism due to their permanent or transient contact with the blood stream or the lymph. A severe contact dermatitis due to endotoxin contaminations of surgical gloves was described in 1984 by Shmunes and Darby. After 8 pyrogenic reactions in altogether 69 patients undergoing heart catheterization, Kure et al. described endotoxin contaminations of extracts of the hospital's surgeon's latex gloves which evoked fever in the rabbit and could be successfully transmitted to cardiac catheters in 1982. Grötsch et al. were able to evoke fever reactions in rabbits with the eluate of gloves containing up to 2560 EU on their inside (Grötsch et al., 1992). Apart from that there are materials that are absorbed by the body, thus taking up any possible inherent pyrogenic materials, such as resorbable sewing materials. According to German legislation, implants are considered parenterals and therefore are drugs. For these products, according to the monograph "Parenterals" of European Pharmacopoeia, pyrogen testing is not required, but is suggested by several authors (Bohner at al., 1994, Grötsch and Eibach, 1990). The AAMI stated in 2001 that products with direct or indirect contact to the circulation system or the lymph or which interact systemically with the body should be tested for pyrogens (AAMI, 2001).

In order to judge a possible contamination, an eluate of the respective material must be either injected into the rabbit or used in the LAL. The alternative of transplanting the questionable

material directly into the rabbit is highly invasive causing possible reactions not associated with pyrogenic contaminations and is therefore questionable in its ethical and scientific implications. The obvious advantage of the IPT over the classical test methods is that the direct contact of the whole blood with the respective device does not require an eluate that allows only pyrogens to be detected which are not elutable for some reason. Additionally, unlike the LAL, the IPT detects all relevant pyrogens, not only endotoxin, in a species-specific manner. The basic principle consists of bringing the diluted human whole blood into direct contact with the surface of the material to be tested and incubate the blood for 10-24 hours at 37°C, like it has been established with the original method. Medical devices in this case included e.g. alginate microcapsules as carriers of drugs where the liquid material as well as the end product could be identified as pyrogen-free and an endotoxin control as well as a non-endotoxin spike was retrieved in an interference-free manner when compared to the same spikes in saline (see figure), dialysis membranes, dialysis fluids, bone substitute materials like tricalciumphospates, and implants such as hip joints. When testing filters, several, though not all, filter materials could be recognized in the WBT as pyrogenic or non-pyrogenic in an interference-free, dose-dependent manner.

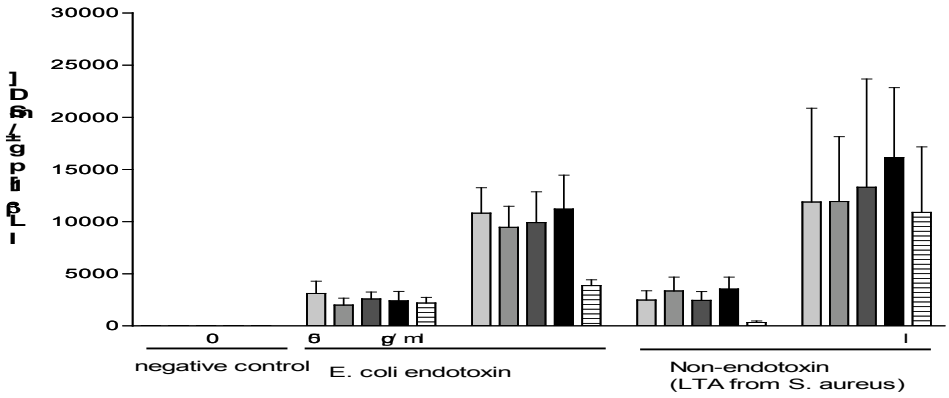


Figure: IL-1□ production of fresh blood of 4 donors upon stimulation with clean alginates and an artificial endotoxin/ non-endotoxin spike

Metallic and plastic surfaces

The testing for the inflammation inducing potential of implant surfaces for the judgment of biocompatibility is a relatively new field. In the early 1980s, it was noted that the monocyte is one of the first cells to arrive at an implant site and displays manifold functions (for review see Anderson, 1984; Ziats 1988). Its specific preference for rough and hydrophobic surfaces which develops after 7 days was termed rugophilia (Rich and Harris, 1981). In 2002, Soskolne et al. documented the adhesion and secretion of TNF-□ of monocytes on titanium surfaces that increases with increasing degrees of roughness; a finding that was later confirmed by Refai et al., 2004 who extended the study of Soskolne to IL-1□ and IL-6 as well. The authors found an enhanced secretion of all three proinflammatory cytokines towards a low LPS dose on the roughest titanium surface with no IL-1□ and IL-6 secretion of unstimulated cells. In contrast, TNF-□ secretion was elevated on this surface, even without LPS stimulus. Another titanium surface that had been polished tended to diminish IL-1□ and IL-6 secretion after LPS stimulation (Refai et al., 2004). The role of cytokine production of the monocytes/macrophages in the early stages of implant insertion are until now poorly understood. The fact that obviously some materials are capable of modulating the cytokine response makes it difficult to distinguish a genuine pyrogenic contamination from an unspecific activation and poses the problem of adequate negative controls. For this purpose, in a master-thesis by Kullmann in 2002, a model was developed for the testing of metallic or plastic surfaces. The study showed that pyrogenic contaminations on surfaces could be reliably removed only when heated for 5h at 300° C with no differences in varying surface

structures (rough or smooth). This applied to titanium, titanium alloy (TiAl6V4) and steel material for implants. The blood was incubated directly in a depyrogenized microtiter plate with contact to the surface to be tested. Artificial contaminations could be recognized in a dose-dependent manner and removed by different washing/heating procedures. With this procedure, safety testing of medical devices with the IPT can be performed, with the possibility of testing the material itself without an eluate and an adequately depyrogenized negative control.

Aneurysm clips made of titanium

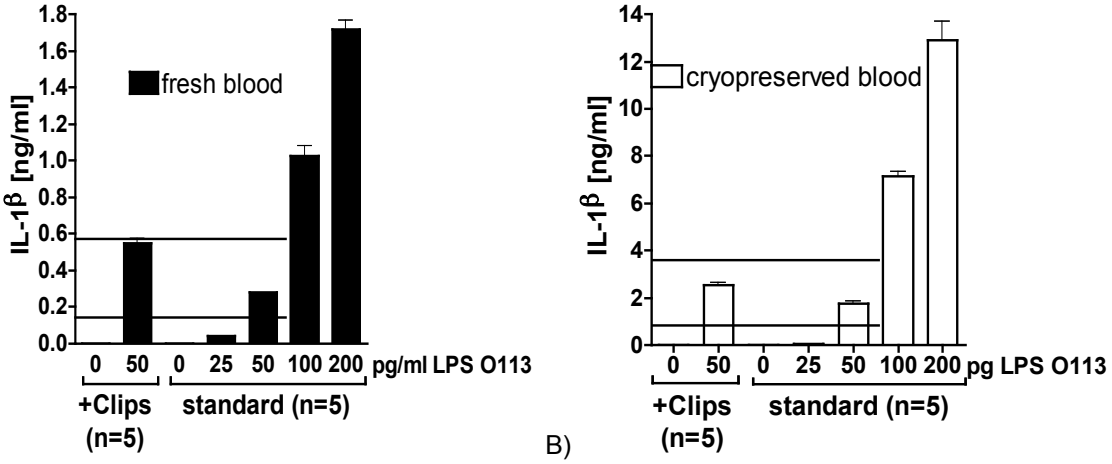


Figure: Comparison of the use of fresh and cryopreserved human blood for pyrogen detection on titanium clips.

5 clips each were incubated with 100 µl saline or LPS (50 pg/ml) diluted in 100 µl incubation medium. Then 100 µl fresh blood (panel A) or cryopreserved blood (panel B) were added and IL-1□ release was measured by ELISA. Data are given as mean ±SD or rather as median. The horizontal lines mark 50% and 200% of the IL-1□ release in response to 50 pg/ml LPS O113 in the absence of a clip.

A) Titanium (aneurysm clips)

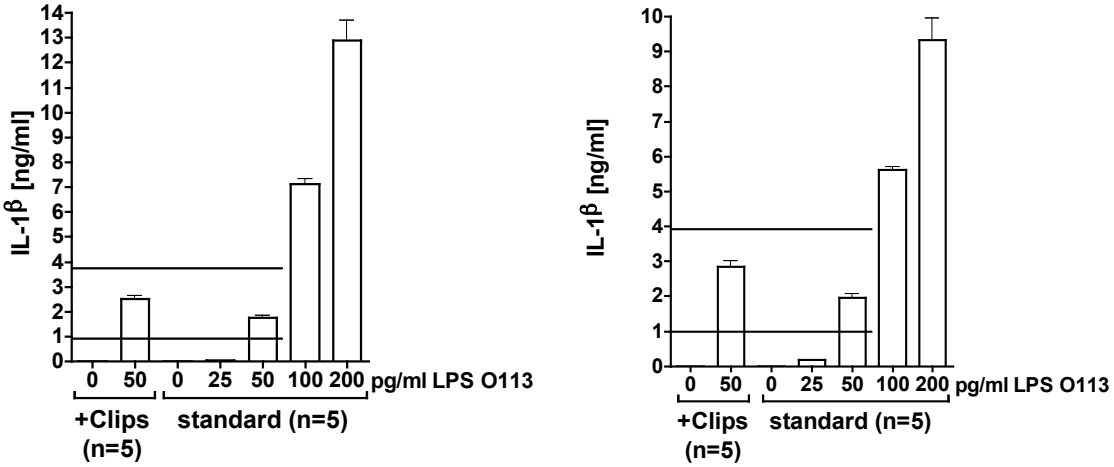


Figure: IPT with aneurysm clips made of titanium

Incubations were performed with cryopreserved blood and IL-1□ was measured by ELISA. Data are given as mean ±SD or rather as median.

material	device	effects on IPT
Chirulen 1020		none
Chirulen 1050		none
metallocene-PE	human blood bags	none
micoporous surface on titanium	chips	none
nickel silver 2.0441	bar stock	increasing
PA6.6.	bar stock	none
PEEK	bar stock	increasing
polyamide coat	ablation catheter	increasing
polyamide coat	balloon catheter	increasing
polyethylene coat	heartwire	none
polypropylene with titanium coating	mesh-implant TiMESH	none
polyurethane Elastollan	pellets	none
polyurethane film	boxes for clips	none (fresh blood)
polyurethane film	boxes for clips	decreasing (cryo blood)
polyurethane foam with polytetrafluoroethylene film	synthetic skin substitute	decreasing
POM	bar stock	increasing
PPSU	bar stock	increasing
stainless steel 1.4021	bar stock	increasing
stainless steel 1.4306	bar stock	increasing
stainless steel 1.4401	bar stock	increasing
stainless steel 1.4542	bar stock	increasing
Titanium	aneurysm clips	none
Titanium 3.7165	bar stock	increasing

Table: Summary of the results of influences of the tested materials on the IPT

Dialysis

In the US, the number of patients receiving dialysis has almost tripled from 63.000 to 170.000 between 1982 and 1992 (Bland, 1995). In the same time, the percentage of dialysis centers that reprocess hemodialyzers for reuse has increased from 19 to 72%. Pyrogenic reactions of hemodialysis (HD) patients at the end of a session were first associated with high bacterial and endotoxin levels by Raij et al., 1973 and Favero et al., 1974.

Since then, possible contaminations could be attributed to the pure water (Klein et al., 1990, Pegues et al., 1992, Kulander et al., 1993, Bambauer et al., 1994) to the machines, especially to areas with low circulation or dead spaces which serve as a reservoir for bacteria (Phillips et al., 1994), the filter materials (Schouten et al., 2000) and bicarbonate concentrates (Pegues, et al., 1992).

In 1993, the AAMI (Association for the Advancement of Medical Instrumentation) released recommendations for the quality of treated water and dialysate which restricted the content of heterotrophic bacteria to 200 and 2000 cfu/ml, respectively. Studies in Germany (Bambauer et al., 1994), Greece (Arvanitidou et al., 1998), the USA (Klein et al., 1990), and Canada (Laurence and Lapierre, 1995) revealed that even these moderate standards are not met, which is even more critical considering that a patient with chronic renal failure receives up to 400l of dialysis fluid a week. For example, in the US, 53% of the treated water did not comply with AAMI standards (Germany 17.8%, Greece 6.4%). According to the study of Bamberger et al., in Germany, 12.2% of the water samples contained 5 EU/ml or more with dialysate contaminations of up to 487 EU/ml. Pseudomonas was found in the water of 73% of the centers (dialysate 90%). Cocci (Micrococci, Staphylococci and Streptococci) were found in the dialysate of 83, 70, and 10% of the centers, respectively, indicating the importance of Gram-positive contaminations. Nakagawa et al., 2002, did an extensive study of the detection of non-endotoxin pyrogens in the whole blood assay and a cell line in comparison to the rabbit. That this might indeed be crucial for judging the pyrogenic load for a dialysis patient was assessed by Marion-Ferey et al., who tested scrapings of bacterial

biofilms in dialysis tubes and found a 20fold higher response in the IPT than in the LAL (Marion-Ferey et al., 2005). Petri et al., 2000, also reported better recognition of Gram-positive and Gram-negative stimuli in the WBT when compared to the LAL.

The passage through dialysis filters of cytokine-inducing substances, not only endotoxins, but exotoxins and peptidoglycans as well, have been demonstrated (Evans and Holmes, 1991, Lonnemann et al., 1992, Urena et al., 1992, Tsuchida et al., 1997). That this indeed can pose a severe threat to patient health has been the result of extensive research over the past years (Schindler et al., 1996, Lonnemann 2000). More than one clinical pyrogenic reaction per year occurred in the US in about 20% of all dialysis centers (Tokars et al., 1991).

A majority of authors found elevated cytokine levels in unstimulated patient blood or an enhanced reactivity of the monocytes of uremic HD patients towards low doses of LPS when compared to a healthy group, on the mRNA level (Pertosa et al., 1993, Schindler et al., 1993, Girndt et al., 1995), on the protein level (Haeffner-Cavaillon et al., 1989, Girndt et al., 1995) as well as recently in vitro employing the whole blood test (Vaslaki et al., 2000, Canaud et al., 2001). Girndt et al. could associate the higher IL-6 production in vitro with elevated in vivo IL-6 serum levels and showed an impeded immune response towards vaccination in these patients. That IL-1□ plays an important role in the pathogenesis of HD related complications were reviewed by Dinarello in 1988. Elevated intracellular IL-1□ in unstimulated patient samples and a reduced reactivity to a high LPS dose (2µg/ml) was found by Blumenstein et al., 1988. Donati et al. showed in 1997 elevated unstimulated plasma levels of IL-1□ and IL-1ra as well as a strongly increased proportion of monocytes expressing the IL-1 receptor (8.7± 1.9% in healthy subjects compared to 31.5 ± 3.5% in HD patients).

The long-term consequences of chronically increased cytokine levels seem to be even more relevant. Even if cytokines are not the cause of amyloidosis, they might aggravate the process. In 1991, Baz et al showed that the use of ultrapure water delays the onset of the carpal tunnel syndrome. The group of Schwalbe showed in 1997 in a retrospective study that the incidence of amyloidosis decreased between 1988 and 1996 along with the disappearance of a water softener known to promote bacterial growth and the introduction of reverse osmosis, a very effective method for purifying water. A connection between other phenomena such as malnutrition, poor immune responses and high incidence of malignant tumors in long-term HD patients and cytokine production has yet to be firmly established.

A possible method for testing dialysis fluids in the WBT used varying percentages of diluents and samples (unpublished results of the developing laboratory). While the classic WBT protocol involves 1 ml of 0.9% saline, this was replaced by 1 ml of pyrogen free water in order to compensate the high electrolyte content of e.g. bicarbonate fluid. For the testing of water, a reverse protocol was employed, using 1 ml of the water to be tested in order to lower the detection limit of the test (usually 0.25 EU/ml) and instead of sample 100 µl of 40% sterile saline (Table).

Basic protocol	Modified protocol for electrolyte solutions	Modified protocol for dialysis water
1000 µl of 0.9% saline	1000 µl of pyrogen-free water	1000 µl dialysis water
100 µl sample	100 µl electrolyte solution	100 µl 40% saline
100 µl blood	100 µl blood	100 µl blood

Table: Possible modification of the WBT standard protocol for the testing of dialysis fluids

Notably, a variant of the WBT, termed AWIPT (see responses to questions 4 & 5), offers the opportunity via endotoxin extraction and accumulation to lower the limit of detection and remove interfering substances, both relevant for dialysis fluids.

Lipidic formulations

Since January 2004, the testing of so-called small volume parenterals (< 15ml) has been made obligatory by European Pharmacopoeia. This concerns many formulations that have not been subjected to pyrogen testing before such as vitamin preparations and steroids. Many of these are applied intramuscularly or subcutaneously and therefore not necessarily have a hydrophilic nature. This poses a completely new challenge to all methods of pyrogen testing, since a lipophilic substance cannot be injected intravenously into the rabbit and will, on the other hand, influence the OD measured in the BET due to the formation of miniscule oil drops. Furthermore, the pyrogenic portion of the LPS, the lipid A (for review see Rietschel et al., 1993) has been reported to be masked by lipoproteins (Emancipator et al., 1992) and lipophilic parenterals (Paulssen and Michaelsen, 1984) in the LAL. Therefore, the WBT procedure was adapted to suit lipophilic substances (Schindler S., submitted). As a first step, interference-free oils such as sesame oil were identified by comparing an LPS dose response curve in these oils with a similar curve done in saline. Surprisingly, many oils (sesame oil, peanut oil, paraffin, miglyol) were interference-free, while others, and especially drug-containing end products, interfered strongly by suppressing the endotoxin stimulus added. Oils that tested interference-free were then used as diluents for interfering end products. It was possible to dilute the interference to non-detectable limits with full recovery of an artificial endotoxin spike. From this minimum valid dilution a possibly detectable endotoxin concentration could be calculated, which was 20 EU/ml for the respective end products. Since these products are applied at a very small volume (1ml per person) a relatively high endotoxin contamination can be tolerated which will predictably not pose a health hazard for the recipient. The established protocol leaves a broad safety margin, especially since the criteria for intravenous drugs were applied (Schindler et al., submitted).

Conclusion: The WBT using fresh as well as cryopreserved blood is a useful and reliable tool for several aspects of pyrogen detection. Not only does it detect a wide spectrum of possible fever-inducing substances, but also its robustness makes it available for such different aspects as the testing of solid substances, (immuno-) toxic drugs, air quality, and biologicals.

MM6/IL-6
Testing of medical devices, immunoglobulins and antibodies has not been performed and additional testing is not foreseen in the developing laboratory in the near future.
Testing of vaccines using a MM6/IL-6 assay has been very limited so far (Carlin and Viitanen, Pharmeuropa Vol 15, no 3, 2003, page 418-423). Results with the multivalent vaccine Infanrix suggest that (some of) the components inhibit IL-6 production by MM6-cells (Infanrix interfered also with the BET assay). Preliminary experiments also indicate that interference should be assessed.

4. There are well-known limitations of the BET (e.g., proteins that interfere with LAL, chemical extracts from medical devices). What are the specific limitations of the proposed test methods? Do certain physicochemical properties interfere with these assays (e.g., pH, osmolarity, protein content) and, if so, what are they and how do they encumber these test methods?

In general, the novel tests are not suitable for drugs that interfere with the cytokine readout. This applies to all drugs that either raise or diminish the signal. In order to judge this positive or negative interference, the Limulus criterion of a 50-200% retrieval was chosen, comparing the spiked sample to the reference control value in saline. The SOPs of all tests include sections on testing for interference. The available data show that interference testing is necessary and acceptance criteria need to be established for the individual pharmaceutical products.

The toxicity and pyrogenicity in man is well known for LPS but only from E. coli. For other endotoxins e.g. from Pseudomonas or Salmonella we know that they can be 1000 times less or 10 times more toxic. About combinations of endotoxins or even combinations of endotoxins and non-endotoxin pyrogens there is very limited experience. Moreover there is some concern about potentiation of the activity of low endotoxin contamination together with a usually non-pyrogenic contamination like glucans or pepdidoglycan (Traub S, von Aulock S, Hartung T and Hermann C. MDP and other muropeptides – direct and synergistic effects on the immune system. J Endotox Res 2005, in press.; Traub S, Kubasch N, Morath S, Kresse M, Hartung T, Schmidt RR and Hermann C. Structural requirements of synthetic muropeptides to synergise with LPS in cytokine induction. J. Biol. Chem. 2004, 279:8694-8700.). The conclusion is that every drug product has to be validated individually and the acceptance criteria have to be defined very carefully (together with statisticians and physicians). The testing for interference is described in all SOPs of the novel tests.

The individual tests differ slightly with regard to the dilution of the test sample (from 1:1 to 1:12), which results in some differences with regard to the limit of detection but also how much they are prone to interference with the products. The WBT tests are extremely robust due to the buffer capacity of human serum (e.g. allowing samples of pH 1 or pH 14 to be tested) and the strong dilution (1:12).

WB/IL-1
A selection of drugs all known to interfere with the rabbit pyrogen test and in part also with the Limulus test, i.e. taxol, was used to test the efficiency of pyrogen detection by WB/IL-1 and the newly developed AWIPT. The AWIPT (Absorb and Wash In vitro Pyrogen Test) is a new development that enables all drugs interfering with the WBT to be tested. Endotoxins are extracted by endotoxin-binding beads, which are after a wash step that takes out the interfering materials, exposed to the WBT. This makes use of the unique property of the WBT to allow endotoxin determination on solid materials. The AWIPT allowed spike recovery at a dilution at least a factor 3 less than in WBT for all drugs except for gentamicin spiked with 25 pg/ml LPS where the recovery was the same in both tests. The safety margin required to exclude relevant pyrogenic Intentional contaminations with lipopolysaccharide were retrieved from the chemotherapeutic agents paclitaxel, cisplatin and liposomal daunorubicin, the antibiotic gentamicin, the antifungal agent liposomal amphotericin B, and the corticosteroid prednisolone at lower dilutions than in the standard in vitro pyrogen test.

In this study, such problematic substances were addressed using macroporous acrylic beads decorated with immobilized human serum albumin (HSA). The material was originally developed as an extracorporeal endotoxin-adsorbing matrix to treat endotoxemia and sepsis. Albumin is a universal carrier of lipophilic substances like fatty acids, bilirubin and hormones in plasma via hydrophobic interactions in a molar ratio of up to 10:1. This test represents a promising new approach to test interfering drugs or drugs containing interfering additives for pyrogenic contaminations, thus improving the safety level of the drugs.

	WBT	AWIPT	WBT	AWIPT
	Dilution at recovery of spike (25 pg/ml LPS)		Dilution at recovery of spike (50 pg/ml LPS)	
paclitaxel	100	10	100	31
gentamicin	31	31	100	31
Cisplatin	100	10	>316	31
prednisolon	> 316	10	>316	31
liposomoal amphotericin B	10	3	31	10
liposomal daunorubicin	>316	31	>316	100

Table: Minimum dilutions of complex therapeutics at recovery of LPS spike.
Drugs were diluted in series and spiked with 25 pg/ml or 50 pg/ml LPS. Recovery of spike, defined as cytokine release at levels between 50% and 200% of those induced by the same concentration of LPS in the absence of the drug, was achieved at the given dilutions.
Additionally, the recognition of LPS as well as LTA can be improved by immobilizing it on a surface. Therefore, the AWIPT cannot only be used to enhance LPS detection, but that of LTA as well.

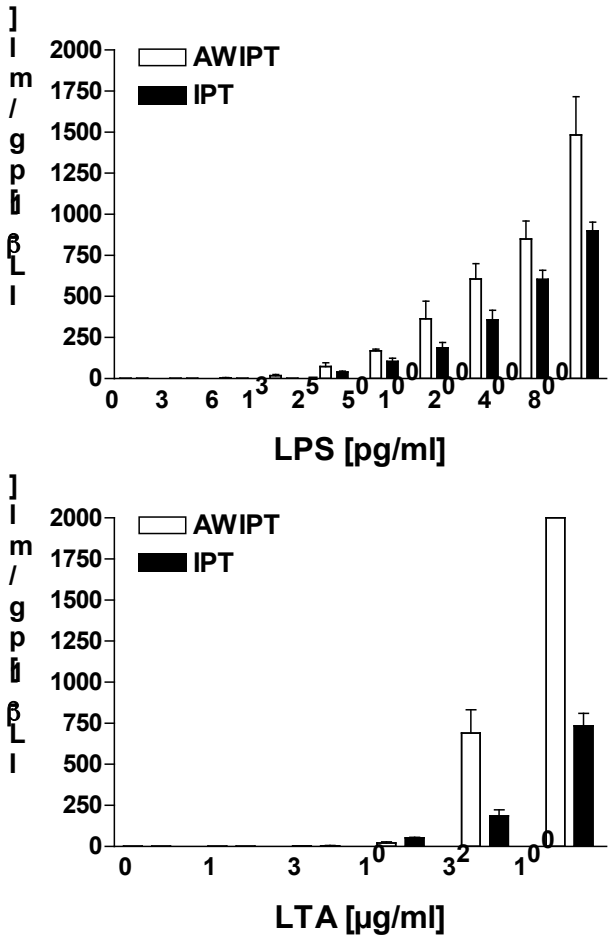


Figure: Comparison of IL-1β response to LPS and LTA in IPT and AWIPT
100 µl samples of 3 to 800 pg/ml LPS or 1 to 100 µg/ml LTA were employed in parallel in IPT and AWIPT. IL-1β release was measured by ELISA. Data represent means ± SD of triplicates from one blood donor.

5. The current bacterial endotoxin test requires standardization of the Limulus amoebocyte lysate (LAL). How would the cellular components of the proposed test methods (i.e., whole blood, PBMCs, cultured monocytoid cell line) be standardized?

WB/IL-1 & WB/IL-6
Work in several hundred blood donors has shown that the threshold of cytokine induction and the levels of cytokine released are sufficiently conserved among individual donors. Extreme reagibilities are very rare and are controlled by the controls employed. Since the donor's individual response curve to endotoxin serves as calibrator in each measurement, any difference in

responsiveness is controlled for. The availability of standardized kit versions including control reference endotoxins further contributes to the standardization of the test.
Further standardization is achieved by cryoconservation of blood as demonstrated in the validation. A blood donation as for transfusion purposes (500ml) by five pooled donors would suffice for up to 25.000 measurements. Shelf lives of more than one year have been established at -80 degree and liquid nitrogen, respectively. Certified cryoblood produced according to GMP and ISO standards is already commercially available.

PBMC/IL-6
The method with PBMC/IL-6 sees 2 types of test: First of all, an investigation for interferences of the test substance with the test system and the readout system has to be performed and then the highest test concentration not showing interferences is determined. The first type of test is to calibrate each donor response on an individual calibration curve produced with the standard endotoxin. The donor response is then defined in terms of "Endotoxin Equivalents" and the product specification is equal to the endotoxin limit concentration (ELC). The standardization goes through multiple donor testing and setting of adequate acceptance criteria. The second type of test (as described in the SOP) is to test against a "clean" reference preparation of the same product. The standardization is again assured by testing of PBMC coming from different donors.

MM6/IL-6
In our opinion the proper procedure to standardize the MM6/IL-6 is equivalent with the procedure described for PBMC/IL-6. Interference of a drug product with the cells and the readout system should be tested with a number of different batches of the same product (known to be pyrogen free). The highest test concentration not showing interference (or an acceptable level of interference) is determined. This concentration should be applied for testing suspect products. The suspect product should always be tested against a “clean” standard reference preparation of the same product.

6. The prediction model described in the BRDs is based on a pyrogen threshold concentration of 0.5 EU/mL. While this level of detection would indeed suffice for many parenteral drugs and medical devices, the endotoxin limit set by the U.S. Food and Drug Administration for intrathecal drugs and devices that contact cerebrospinal fluid is 0.06 EU/mL. Do you have data to support the use of the proposed test methods for discriminating an endotoxin threshold lower than 0.5 EU/mL?

The immune system of all organisms reacts extremely sensitively to bacteria recognizing conserved structures often termed endotoxins. Here, man, rabbit and horseshoe crab do not differ very much. The thresholds of reaction are remarkably conserved in the low picogramme range or about 0.1 to 1 bacteria per immune cell. The precise set-up of the test is determining the limit of detection. For the purpose of predicting the rabbit response, tests were adjusted to a threshold of 0.5 EU/ml. Given a routine dilution of the samples of up to 1:12 and measurable signals also at endotoxin concentrations lower than the 0.5 EU threshold demonstrates that the sensitivity of the systems has not been fully exploited.

WB/IL-1
The test has been modified to include adsorption of endotoxin to beads (termed AWIPT, see above). This offers the possibility, by concentrating the LPS on its surface and enhancing the reactivity of the monocytes, to detect as little as 0.0001 EU/ml (see Figure)

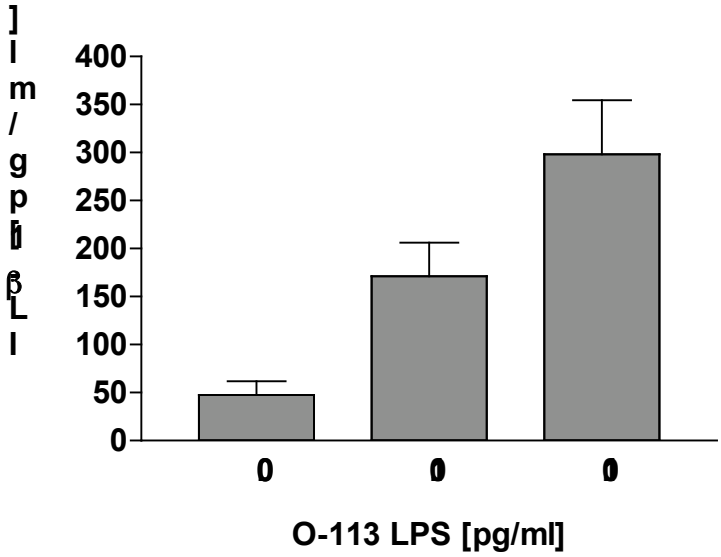


Figure: Limit of detection in the AWIPT

MM6/IL-6
During the development of the assay it was chosen to settle for a threshold concentration of 0.5 EU/ml. However, the sensitivity of the MM6/IL-6 is as low as 0.1 EU/ml in most experiments. Aiming at a threshold level of 0.06 EU/ml will challenge the assay.

PBMC/IL-6
The sensitivity against standard endotoxin of the test with PBMC/IL-6 is comparable to the BET. The detection limit is about 0.01 EU/ml.

Comments and suggestions relevant to all BRDs

* ***A list of definitions would be useful.***

Such a list will be included.

* ***There appear to be different designations for the cells/tests in the BRDs, in document Stp-HPTVv04, Comparison and Validation of Novel Pyrogen Tests Based on the Human Fever Reaction. Trial Plan, and the file names and nomenclature of the tests in the SOPs. The cell/test nomenclatures should be standardized to avoid confusion. Examples of various designations used are: PBMC-IL6; PBMC; WB-IL6; WB/IL-6; WB-IL1; cryo WB-IL1; WB-CRYO/IL-1; MM6/IL6; MM6.***

This will be amended as far as it is not part of historic documents, where tests were termed differently.

* ***Sections 4.1 and 4.6 comment on theoretical assumptions of sensitivity and specificity, and cite reference [10]. There is no reference [10] in the BRDs; the citation, if relevant, should be provided in the format used for the BRDs.***

This has been corrected.

* **The BRDs mention that the SOPs for the different tests are in Section 13.1. However, Section 13.1 only references the SOP; copies of the SOPs are included in Section 15 of each BRD.**

The BRDs have been revised and the method protocols and trial plans are now included in Appendix A.

* ***In all the submitted studies, the accuracy of the test is being measured using bacterial endotoxin (LPS) in all test samples; presumably only the vehicle is changed. However, one of the claimed advantages for this test over the BET is that it is capable of detecting non-LPS pyrogens, whereas the BET cannot. There are no test results from the non-LPS pyrogens referred to in the Rationale. In the absence of additional data on other pyrogens, it cannot necessarily be assumed that the tests would be relevant for non-LPS pyrogens.***

See 2)

* ***The conclusion in Section 6.4 that the test is applicable to "most classes of medicinal products> needs expansion and clarification. First, the product classes to which the test is not applicable should be identified; there appears to be no information in the BRD at this point. Secondly, the statement should be clarified to state that the test is applicable to the detection of LPS. No evidence has been presented with respect to other classes of pyrogens.***

Rabbit pyrogen testing as the reference method has been substituted by large extent with the BET, which is a mere endotoxin test. The remaining rabbit testing is due to interference of test materials with the BET and not due to its limitations to endotoxin. A novel substitute for the rabbit test should be evaluated on the same basis. Since non-endotoxin pyrogens have not been internationally agreed and made available as reference materials, a formal validation is not possible. The supportive information that the in vitro pyrogen tests cover in fact some of the presumed non-endotoxin pyrogens represent a characteristic in favour of these tests compared to the BET.

* ***There is a comment in the BRDs (Section 4.1) regarding historical data from rabbit tests, yet the sensitivity, specificity, and concordance values presented are not related to the in vivo data (i.e., Section 6.2, Section 6.3). It is not clear how these values could be generated without relating to rabbit or human test results on the same samples.***

On the basis of the determined rabbit fever threshold (Hoffmann S, Luderitz-Puchel U, Montag-Lessing U and Hartung T. Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modelling, J. Endotoxin Res. 2005, 11:25-31), it was possible to model the performance of these rabbits, when testing the samples of the validation study assuming no additional interference of the samples. The sensitivity of the rabbit pyrogen test is 57.9% and the theoretical specificity of the rabbit pyrogen test is 88.3%. Unfortunately these numbers have not been correctly reported in the BRD, in sections 4.2 and 4.6., where the last sentence reads " .. sensitivity of the rabbit pyrogen test is 75.04% and the theoretical specificity of the rabbit pyrogen test is 95.80%.

* ***There appears to be little relationship among the articles cited in each BRD, the list of references in Section 12 of each BRD, the articles included in Section 15 of each BRD, and the articles supplied in electronic format. This could be a problem if a reviewer wanted to read a referenced article. The reference citations in the text, the list of references in Section 12, and the copies of the references in Section 15 need to be coordinated. These discrepancies are presented in different levels of detail in the following BRD assessments.***

The references have been corrected and hardcopies of a number of publications are included in Appendix B.

* ***There are statements in Section 9.3 of all the BRDs that compare the performances of different tests, but data are not presented. The data are in the referenced Hoffmann publication, which was submitted on a CD file, but not in hard copy. It would assist in the assessment of these assays if the Hoffmann article could be appended to the individual reports, and if summary performance tables were included in each report to support the brief verbal description of test performance that appears in Section 9.3.***

Added as requested.

* ***It would be helpful if the articles in Section 15 were in alphabetical order.***

The hardcopies of a number of articles are in Appendix B.

* **It would be helpful to the reviewers to have a table comparing the strengths and weaknesses, if any, of the assays.**

See table 9.3.2 in Section 9 of the BRDs

* ***Some, if not all, of these tests are patented. The patented tests and procedures should be identified. The Sections on test method transferability (Section 11.1) and cost (Section 11.3) should address the availability, licensing fees and licensing agreements, if any, of these tests.***

This is now mentioned in the individual BRDs.

Comments on the individual BRDs

Comments relevant to the Human PBMC/IL-6 In Vitro Pyrogen Test

* *Section 1.1.2. The results of the cited FDA peer review have not been summarized or provided.*

The following information is now included: The PBMC/IL-6 test developed by Novartis and Baxter Healthcare has been subjected to a rigorous peer-review by the US FDA and approved as an end-product release test (New Drug Application Number 16-267/S-037 approved on April 24, 2002).

* *Section 2.1. The file reference in Section 2.1 should name the specific file because there are two PBMC SOP files, and there is no information provided as to which SOP file is referred to in the BRD.*

This has been amended in Section 2.1 and Section 13 (catch-up validation of PBMC CRYO IL/6) and Appendix A includes both method protocols.

* *Section 2.3. Define LAL*

LAL = Limulus amoebocyte lysate

Please note that the term "LAL" has been replaced in the 5 BRDs with the more general term BET (= Bacterial Endotoxin test), which is based on the use of Limulus amboecyte lysate.

* *Section 2.4. The data presented here were obtained using a Novartis-developed IL-6 ELISA assay, and this section states that any commercial IL-6 ELISA kit will have to be validated for this pyrogen test. Unless the Novartis assay will be publicly available, non-Novartis users (who, presumably, the test is designed for) will have to go through a separate validation of this assay.*

Any human IL-6 ELISA can be used provided International Standard (IS) for IL-6 (or an IL-6 standard calibrated against the IS) is used as the assay calibrant. In addition, it would be necessary to demonstrate that the requirements of test controls are met and no interference with the test substances occurs.

For the two other methods with IL-6 release as an endpoint used either the Novartis IL-6 ELISA (WB/IL-6) or a commercial kit (MM6/IL-6).

* *Table 3.3.1. Define > "> notional ELC> "> .*

notional ELC = endotoxin limit concentration set by the European Pharmacopoeia monograph (or other guidelines) for a given product. The term is explained in each of the five BRDs.

* *Section 12. There are a number of discrepancies among the cited articles, the bibliography, and the provided references. Many of the publications listed here do not correspond with those cited in the submission or those included as hard copies in Section 15.*

This was resolved. Cited articles are listed in Section 12, articles in bold are attached as hardcopies in newly created Appendix B. Appendix B includes the same set of publications for the 5 BRDs.

Comments Relevant to the Human Whole Blood/IL-1 In Vitro Pyrogen Test

* There are no comments specific to this test method. There are reference and citation problems similar to those identified in Sections 2.6.2 and 2.6.6.

This was resolved. Cited articles are listed in Section 12, articles in bold are attached as hardcopies in newly created Appendix B. Appendix B includes the same set of publications for the 5 BRDs.

Comments relevant to the Human Whole Blood/IL-1 In Vitro Pyrogen Test: Application of cryopreserved human whole blood

* *Although this is a test method that uses cryopreserved blood, the SOP in Section 15 does not address cryopreservation.*

Unfortunately, there was a mistake with the first submission, the correct protocol is now given in Appendix A of BRD CRYO WB/IL-1. In addition, the article Schindler et al, 2004 in appendix B deals with cryopreservation.

* *Section 2.5. An abbreviated validation study was performed. The validation study of this method appears to consist only of a comparison of the results from using cryopreserved blood with the results from the same test (WB/IL-1) using fresh blood. Therefore, the statements on (intra-laboratory) reproducibility should be removed from Section 3.1 and Section 5 because Section 2.5 and the data in Section 5 indicate that reproducibility was not examined.*

Data of on intralaboratory reproducibility are included in Appendix D of the BRD CRYO WB/IL-1.

* *There are reference and citation problems similar to those identified in Sections 2.6.2 and 2.6.6.*

This was resolved. Cited articles are listed in Section 12, articles in bold are attached as hardcopies in newly created Appendix B. Appendix B includes the same set of publications for the 5 BRDs.

Comments Relevant to the Human Whole Blood/IL-6 In Vitro Pyrogen Test

* *Section 2.4. The data presented here were obtained using a Novartis-developed IL-6 ELISA assay, and this section states that any commercial IL-6 ELISA kit will have to be validated for this pyrogen test. Unless the Novartis assay will be publicly available, non-Novartis users (who, presumably, the test is designed for) will have to go through a separate validation of this assay.*

Any human IL-6 ELISA can be used provided International Standard (IS) for IL-6 (or an IL-6 standard calibrated against the IS) is used as the assay calibrant. In addition, it would be necessary to demonstrate that the requirements of test controls are met and no interference with the test substances occurs. For the two other methods with IL-6 release as an endpoint used either the Novartis IL-6 ELISA (PBMC/IL-6) or a commercial kit (MM6/IL-6; see below).

* *There are reference and citation problems similar to those identified in Sections 2.6.2 and 2.6.6.*

This was resolved. Cited articles are listed in Section 12, articles in bold are attached as hardcopies in newly created Appendix B. Appendix B includes the same set of publications for the 5 BRDs.

Comments Relevant to An Alternative In Vitro Pyrogen Test Using the Human Monocytoid Cell Line MONO MAC-6 (MM6)

* *Section 2.6. This suggests that only the Novartis IL-6 ELISA assay is usable because other ELISAs were not repeatable, and therefore could not be used. This aspect should be expanded upon because it suggests that users of the MM6 test will be limited in the ELISA preparations they can use or, alternatively, have to validate the test using other ELISAs.*

The initial protocol allowed the use of various kinds of IL-6 ELISAs, however, due to their sub-optimal repeatability their use was restricted to the two IL-6 ELISAs now indicated in the protocol (the in-house Novartis IL-6 ELISA and the CLB Human IL-6 ELISA kit). Both IL-6 ELISAs use the same monoclonal antibody for IL-6 detection. It should be noted that these ELISAs may be substituted with other validated IL-6 specific ELISAs, in which the International Standard (IS) for IL-6 (or an IL-6 standard calibrated against the IS) is used as the assay calibrant.

* *Section 2.4. The cell line that was used for this validation study is available only from a particular laboratory at the University of Munich, Germany, which will supply the cells to all who request them, or from a Master Cell Bank and a Working Cell Bank at the NIBSC (in Germany) (see also, SOP Section 6.1). Another source of cells is the German DSMZ. It is stated here that the performance of the DSMZ-source cells have not been compared with the NIBSC cells used in this validation study, and there is no mention of whether the performance of the NIBSC cells was compared to the Univ. of Munich cells. All cell sources are listed under the heading of > "> Proprietary Components.> "> The BRD should address whether these cells are proprietary, and if there will be a one-time cost or licensing fee, or a licensing agreement, before they can be shipped to a testing laboratory. The sponsor should ensure that the various cell lines are interchangeable in the assay; otherwise, there will be only a single source for the cells. Alternatively, the validated cell line should be more widely distributed so that there would be less likelihood of loss, and so that users outside of Europe will have more easy access to them. Therefore, it may be reasonable to recommend that the sponsor ensure the continued availability of a cell line or lines whose performance of the test is well documented before ICCVAM invests time and resources on the evaluation of this test.*

The performance of the cells obtained from NIBSC (UK) was not compared to cells directly obtained from the University of Munich, Germany or the DSMZ.

The MM6 cell line was established by Prof. H.W.L. Ziegler-Heitbrock, Institute for Immunology, Universtiy of Munich, Munich, Germany. The cell line can be obtained for research purposes only from Prof. Ziegler-Heitbrock or from the German Collection of Microorganisms and Cell Cultures (DSMZ) Braunschweig, Germany. The conditions for licensing of the cell line are to be negotiated individually with Prof.Ziegler-Heitbrock.

* *Section 12. There were a number of discrepancies among the cited articles, the bibliography, and the provided references.*

This was resolved. Cited articles are listed in Section 12, articles in bold are attached as hardcopies in newly created Appendix B. Appendix B includes the same set of publications for the 5 BRDs.

* *The Section 12 reference list contains 27 references; 15 of which are in Section 15. There are 13 articles included in Section 15 that are not listed in Section 12; one of these is also on the CD file. There are 6 articles on the CD file, one of which is also listed in Section 12, and another of which is included as hard copy in Section 15.*

see above

[This Page Intentionally Left Blank]

Appendix C

Additional Information Requested by the Panel

C1 **ESAC Statement on the Validity of *In Vitro* Pyrogen Tests C-3**
C2 **Press Release: "Fewer Tests on Animals and Safer Drugs: New EU Tests Save 200,000 Rabbits per Year" C-9**
C3 **ECVAM Replies to Questions of ICCVAM Pyrogenicity Peer Review Panel C-15**
C4 **Rationale for the Selection of the 10 Substances Tested in the Validation/Catch-Up Validation Study of *In Vitro* Assays for Pyrogen Testing C-21**
C5 **Comparison and Validation of Novel Pyrogen Tests Based on the Human Fever Reaction: Trial Data Report C-25**
C6 **List of Drugs for the Catch-Up Validation Study C-41**
C7 **Analytical Procedure to Identify and Eliminate Outlying Observations C-45**

[This Page Intentionally Left Blank]

Appendix C1

ESAC Statement on the Validity of *In Vitro* Pyrogen Tests

[This Page Intentionally Left Blank]

EUROPEAN COMMISSION
DIRECTORATE GENERAL JRC
JOINT RESEARCH CENTRE
Institute for Health and Consumer Protection
European Centre for the Validation of Alternative Methods (ECVAM)

STATEMENT ON THE VALIDITY OF IN-VITRO PYROGEN TESTS

At its 24th meeting, held on 20-21 March 2006 at the European Centre for the validation of alternative methods (ECVAM), Ispra, Italy, the non-Commission members of the ECVAM Scientific Advisory Committee (ESAC)[1] unanimously endorsed the following statement:

Following a review of scientific reports and peer reviewed publications on the following range of in-vitro pyrogen tests:

1. Human Whole Blood IL-1,
2. Human Whole Blood IL-6,
3. PBMC IL-6,
4. MM6 IL-6, and
5. Human Cryopreserved Whole Blood IL-1,

it is concluded that these tests have been scientifically validated for the detection of pyrogenicity mediated by Gram-negative endotoxins, and quantification of this pyrogen, in materials currently evaluated and characterized by rabbit pyrogen tests.

These methods have the potential to satisfy regulatory requirements for the detection and quantification of these pyrogens in these materials subject to product-specific validation.

The test methods have the capacity of detecting pyrogenicity produced by a wider range of pyrogens, but the evidence compiled for, and considered within this peer review and validation process, is not sufficient to state that full scientific validation of this wider domain of applicability has been demonstrated and confirmed.

Thus, the above test methods can currently be considered as full replacements for the evaluation of materials or products where the objective is to identify and evaluate pyrogenicity produced by Gram-negative endotoxins, but not for other pyrogens.

This endorsement takes account of the dossiers prepared for peer review; the views of independent experts who evaluated the dossiers against defined validation criteria; supplementary submissions made by the Management Team; and the considered view of the Peer Review Panel appointed to oversee the process.

Thomas Hartung
Head of Unit
ECVAM
Institute for Health & Consumer Protection
Joint Research Centre
European Commission
Ispra

21 March 2006

1. The ESAC was established by the European Commission, and is composed of nominees from the EU Members States, industry, academia and animal welfare, together with representatives of the relevant Commission services.

   This statement was endorsed by the following Members of the ESAC:

   Prof Helmut Tritthart (Austria)
   Dr Dagmar Jírová (Czech Republic)
   Prof Elisabeth Knudsen (Denmark)
   Dr Timo Ylikomi (Finland)
   Prof André Guillouzo (France)
   Dr Manfred Liebsch (Germany)
   Dr Efstathios Nikolaidis (Greece)
   Dr Katalin Horvath (Hungary)
   Prof Michael Ryan (Ireland)
   Dr Annalaura Stammati (Italy)
   Dr Mykolas Maurica (Lithuania)
   Prof Eric Tschirhart (Luxembourg)
   Dr Jan van der Valk (The Netherlands)
   Dr Dariusz Sladowski (Poland)
   Prof Milan Pogačnik (Slovenia)
   Dr Argelia Castaño (Spain)
   Dr Patric Amcoff (Sweden)
   Dr Jon Richmond (UK)
   Dr Odile de Silva (COLIPA)
   Dr Julia Fentem (ECETOC)
   Dr Nathalie Alépée (EFPIA)
   Prof Robert Combes (ESTIV)
   Dr Maggy Jennings (Eurogroup for Animal Welfare)
   Mr Roman Kolar (Eurogroup for Animal Welfare)

   The following Commission Services and Observer Organisations were involved in the consultation process, but not in the endorsement process itself.

   Mr Thomas Hartung (ECVAM; chairman)
   Mr Jens Linge (ECVAM; ESAC secretary)
   Mr Juan Riego Sintes (ECB)
   Ms Beatrice Lucaroni (DG Research, Unit F.5)
   Mr Sylvain Bintein (DG Environment, Unit C.3)
   Mr Sigfried Breier (DG Enterprise, Unit F.3)
   Prof Dr Constantin Mircioiu (Romania)
   Dr William Stokes (NICEATM, USA)
   Prof Dr Vera Rogiers (ECOPA)

Annex

The novel pyrogen tests are based on the human fever reaction. Monocytoid cells, either primary from human blood or as propagated cell lines, detect pyrogens of different chemical nature and respond by the release of inflammatory mediators such as cytokines. Since lipopolysaccharides from Gram-negative bacteria are the only type of proven pyrogen, for which an International reference material is available, the tests were standardised to detect the presence of significantly less than 0.5 Endotoxin Units of this preparation, which is considered to be the threshold level for fever induction in the most sensitive rabbit species according to pharmacopoeia test procedures.
The five tests which were sufficiently reproducible and exceeded the rabbit test with regard to sensitivity and specificity for the detection of lipopolysaccharide spiked samples, differ with regard to cell source and preparation, cryopreservation and cytokine measured. The tests have been described elsewhere (1-4). The concept of the validation study (5) and the international validation studies are available (6-7).

1. Poole, S., Thorpe, R., Meager, A., Hubbard, A.R., Gearing, A.J. (1988) Detection of pyrogen by cytokine release. Lancet 8577, 130.

2. Taktak, Y.S., Selkirk, S., Bristow, A.F., Carpenter, A., Ball, C., Rafferty, B., Poole, S. (1991) Assay of pyrogens by interleukin-6 release from monocytic cell lines. J. Pharm. Pharmacol. 43, 578.

3. Hartung, T., Wendel, A. (1996) Detection of pyrogens using human whole blood. In Vitro Toxicol. 9, 353.

4. Schindler S, Asmus S, von Aulock S, Wendel A, Hartung T and Fennrich S. (2004) Cryopreservation of human whole blood for pyrogenicity testing. J. Immunol. Meth. 294, 89-100.

5. Hartung, T., Aaberge, I., Berthold, S., Carlin, G., Charton, E., Coecke, S., Fennrich, S., Fischer, M., Gommer, M., Halder, M., Haslov, K., Jahnke, M., Montag-Lessing, T., Poole, S., Schechtman, L., Wendel, A., Werner-Felmayer, G. (2001) Novel pyrogen tests based on the human fever reaction. The report and recommendations of ECVAM Workshop 43. European Centre for the Validation of Alternative Methods. Altern. Lab. Anim. 29, 99.

6. Hoffmann S, Peterbauer A, Schindler S, Fennrich S, Poole S. Mistry Y, Montag-Lessing T, Spreitzer I, Loschner B, vam Aalderen M, Bos R, Gommer M, Nibbeling R, Werner-Felmayer G, Loitzl P, Jungi T, Brcic M, Brugger P, Frey E, Bowe G, Casado J, Coecke S, de Lange J, Mogster B, Naess LM, Aaberge IS, Wendel A and Hartung T. (2005) International validation of novel pyrogen tests based on the human fever reaction. J. Immunol. Meth. 298, 161-173.

7. Schindler S, Spreitzer I, Loschner, Hoffmann S, Hennes K, Halder M, Brügger P, Frey E, Hartung T and Montag T. (2006) International validation of pyrogen tests based on cryopreserved human primary blood cells. J. Immunol. Meth. 316, 42-51.

[This Page Intentionally Left Blank]

Appendix C2

Press Release: "Fewer Tests on Animals and Safer Drugs: New EU Tests Save 200,000 Rabbits per Year"

[This Page Intentionally Left Blank]

IP

Brussels, 12 May 2003

Fewer tests on animals and safer drugs: new EU tests save 200,000 rabbits per year

New, groundbreaking methods of drug testing to replace animals with safe alternatives, saving up to 200,000 rabbits per year, were unveiled today in Brussels by European Research Commissioner Philippe Busquin. The set of six tests detects potential fever-causing agents (pyrogens) in drugs, by using human blood cells instead of rabbits. The new tests have been developed by a EU-supported research team, involving national control laboratories, test developers, and companies. The tests are being validated by the Commission. They are already being used in over 200 laboratories across the world. Thanks to these alternative methods rabbits will no longer be needed to test the presence of pyrogens in parenteral (non oral) drugs.

"The use of animals to test drugs is unfortunately necessary to safeguard human health," said European Research Commissioner Philippe Busquin. "But we can reduce, replace and refine animal testing, with EU-sponsored research leading the way at world level. The EU's validation of these new testing methods will encourage their broad take-up by industry, ensure drug safety and quality, and reduce the use of animal research. This is an example of the European Research Area in action, developing an environment in which scientific results can be rapidly exploited and transformed into products and processes that improve quality of life, increase competitiveness and benefit animal welfare."

The safety and potency of commercially available medicines and vaccines must be guaranteed. Innovative research, funded and validated by the Commission, aims to replace existing animal-based test methods for fever-causing agents (pyrogens) in parenteral drugs with a new generation of in vitro tests that are more accurate, quicker and more cost-effective.

Blood cells replace rabbits

Understanding of human immunology has advanced rapidly in the past 20 years. Work on human fever reaction and development of test systems for fever mediator molecules, combined with improved cell biology techniques, now enables the innovative use of human cells as biosensors for pyrogens (fever-causing agents). The EU study[1] set out to compare and harmonise six in vitro assays to develop a "state-of-the-art" method for inclusion into the European Pharmacopoeia - which sets the requirements for the quality control of drugs in Europe - thus improving consumer safety.

The EU role

[1] Cell factory project: *Comparison and validation of novel pyrogen tests based on the human fever reaction, with a view to the ultimate replacement of the rabbit pyrogen test and the Limulus assay* (QLK3-1999-00811)

The research project funded by the Commission under the EU Fifth Research Framework Programme (1998-2002) brought together the best teams from academia, industry and regulatory bodies. The Commission's Joint Research Centre (the "ECVAM" facility, or "European Centre for Validation of Alternative Methods") played a major role in the project through provision of scientific and technical advice on the design of the validation study, application of good laboratory practice procedures and distribution and coding of test material.

Industry and regulators jump on board

Interest from both regulatory authorities and industry is very high, with many contributions coming from outside the project consortium that included national control laboratories, test developers, a major pharmaceutical company and a producer of diagnostic kits. For example, the European Pharmacopoeia has set up an international expert group to draft a general method on these new tests. In fact, the tests are already in use in about 200 laboratories worldwide, with great success.

Further take-up and new applications

The Commission will take responsibility for further application of this multidisciplinary, international validation study, including an intended patent. This will encourage successful transfer of the tests and help open new fields for pyrogen testing, such as cellular therapies, medical devices and pollution control in the work place.

Reducing, replacing or refining animal experimentation

Drug quality control is a trans-national matter, which is standardised and regulated in Europe at EU level, thus requiring international collaborative efforts. The European Commission ensures full support for applications to reduce, replace or refine animal experimentation as required by the 1986 Council Directive[2]. This aim is echoed by the European Pharmacopoeia. The "Three Rs" provide a strategy to minimise animal use, without compromising the quality of the scientific work being done.

ECVAM's role is to co-ordinate international validation studies, act as a focal point for the exchange of information, to set up and maintain a database on alternative methods, and to promote dialogue among legislators.

Background: pyrogen and non-oral drugs

Parenteral drugs are commonly employed throughout Europe for treating a variety of illnesses. Ensuring the safety of such widely used drugs requires strict monitoring and control against any possible pyrogenic contamination on a batch-by-batch basis. The most important pyrogen is endotoxin, a constituent of the cell wall of gram-negative bacteria that can generate endogenous fever mediators by white blood cells, particularly monocytes and macrophages.

Rabbits or…

In the rabbit pyrogen test, the test substance is injected into rabbits and any subsequent change in body temperature recorded. A significant rise in temperature indicates the presence of pyrogens. While it has served drug safety control for more than 50 years, it fails for important new therapies such as cellular products or species-specific agents.

[2] *Novel in-vitro testing as alternatives to animal testing*; Council Directive 86/609/EEC

… horseshoe crabs?

Until now, the only in vitro alternative available is the LAL test, based on coagulation of blood from the horseshoe crab (Limulus polyphemus). However the LAL test detects only one class of pyrogens – endotoxins from gram-negative bacteria – leaving patients at risk from "non-endotoxin" pyrogens such as gram-positive toxins, viruses and fungi. It is also subject to interference by various non-pyrogenic substances. And, as it is based on the defence system of an arthropod, it cannot provide results perfectly relevant to humans.

No – human blood cells!

Six alternative cellular assays have therefore been developed to replace the animal rabbit pyrogen test and close the safety gap presented by use of the LAL test in controlling parenterals. All these test systems are based upon the response of human leukocytes (principally monocytes), which release inflammatory mediators (endogenous pyrogens) in response to pyrogenic contamination (exogenous pyrogens).

Quicker, more accurate and more effective

The new tests have several advantages compared with the rabbit test: they are less laborious, cheaper and more sensitive. Results of the validation study suggest that testing on animals can be completely replaced. In contrast to the LAL, the new assays are not restricted to endotoxins from gram-negative bacteria but detect all classes of pyrogens and reflect the potency of different endotoxins in mammals, without suffering interference from endotoxin-binding components in blood products. A commercial kit version for one of the assays has already been developed and standardised, and pre-tested cryopreserved (frozen) blood as a versatile test reagent containing the blood cells as biosensors is under development.

For further information please visit:

http://ecvam.jrc.it/index.htm

http://europa.eu.int/comm/research/quality-of-life/cell-factory/volume1/projects/qlk3-1999-00811_en.html

[This Page Intentionally Left Blank]

Appendix C3

ECVAM Replies to Questions of ICCVAM Pyrogenicity Peer Review Panel

[This Page Intentionally Left Blank]

1 February 2007

ECVAM replies to questions of ICCVAM Pyrogenicity Peer Review Panel

1. Availability of ESAC Peer Review Report

Since we are creating a precedent in making ESAC peer-reviews public, a discussion within ESAC is required, especially since a number of external experts have been involved, who have not been asked. Thus, we are unfortunately not able to make this available at this stage of the process.

2. Lot numbers

a) e-mail of David Allen on 10/01/2007
replied on 12/01/2007 with list of drugs as PDF attached

VALIDATION STUDY: LIST OF DRUGS

Product	Manufacturer	Lot
Alkohol-Konzentrat 95%	B. Braun	2465Z01
Beloc i.v.	Astra Zeneca	DA419A1
Binotal 0,5g	Grünenthal	117EL2
Fenistil	Novartis	21402
MCP Hexal	Hexal	21JX22
Orasthin	Hoechst	W015
Sostril	Glaxo Wellcome	1L585B
Traubenzuckerlösung 5% Eifelfango	Eifelfango	1162

Alkohol-Konzentrat 95% = aethanol 95%
Traubenzuxkerloesung 5% Eifelfango = 5% glucose solution

b) e-mail of David Allen on 12/01/2007
Yes, individual lots were tested in all methods during the validation/catch-up validation study

However, some of the lots used in the validation study were no longer available for the catch-up validation study and one product (Orasthin) was no longer on the market. It was replaced with a product (Syntocinon) containing the same active ingredient. Please find attached the pdf file "List of drugs catch-up validation" and the table below highlighting differences in lot numbers and products.

VALIDATION STUDY (CATCH UP): LIST OF DRUGS

Product	Manufacturer	Lot
Alkohol-Konzentrat 95%	B. Braun	2465Z01
Beloc i.v.	Astra Zeneca	DA419A1
Binotal 0,5g	Grünenthal	117EL2
Fenistil	Novartis	26803
MCP Hexal	Hexal	21JX22
Orasthin	Hoechst	not available
Sostril	Glaxo Wellcome	3H01N
Syntocinon 3 I.E.	Novartis	S00400
Traubenzuckerlösung 5% Eifelfango	Eifelfango	3132

Alkohol-Konzentrat 95% = aethanol 95%
Traubenzuxkerloesung 5% Eifelfango = 5% glucose solution

3. GLP concordance

e-mails of David Allen on 9/01/2007 and on 12/01/2007 (question 1)

a) In vitro data

The initial validation study has been carried out to large extent in laboratories such as National Control laboratories, which do not operate under GLP. It was, however, agreed to comply with the requirements of GLP, especially with regard to the creation and management of SOPs. The partner laboratories have received presentations on the requirements. No auditing was done but various quality checks and blinding mainly under the responsibility of ECVAM were included.
In the catch-up validation, two GLP laboratories and two National Control laboratories participated.

Raw data: In both studies the laboratories were asked to transfer the readings into the excel sheets provided by the biostatistician. This was mostly done by directly inserting the ASCII files created by the plate reader. However, reader printouts are available and can be provided on request.

b) In vivo data and reference to Section 4.4 in the ECVAM BRDs
Indeed it should read here "not applicable" as stated in the WB/IL-6 BRD, since the RBT was not performed during the validation study. As indicated in 4.1 the data used were provided by the Paul-Ehrlich-Institute (PEI; www.pei.de), which is the German Federal Agency for Sera and Vaccines (competent authority) and conducts the RBT according to the European Pharmacopoeia. For further information on the quality assurance established at the PEI please contact Dr Thomas Montag (e-mail: month@pei.de).

This should also be corrected in the main document 4.4 In vivo data quality.

4. Data analysis
e-mail of David Allen 12/01/2007 – question 2

The same data analysis was applied in both studies. The first paragraph in Section 5.3 reads *A generally applicable analytical procedure was employed. This procedure includes a universal PM as well as quality criteria. First, a two-step procedure consisting of a variance-criterion and an outlier-test was applied. For this, the Dixon's test (Barnett and Lewis, 1984), which is USP approved, was chosen with the significance level of* α*=0.01 and applied to identify and eliminate aberrant data.*
Please find attached to our mail, the document 'Trial data report' of the validation study. It was not included in the submission to ICCVAM, since a lot of the procedures described here are included in the BRD. Related to your question, you will find in Chapter 4.2 the procedure describing the exclusion of data. There, the Coefficient of Variation (CV) was used as a trigger to investigate the replicates of a given control or sample. Excessive variability would severely impair the prediction model, resulting mainly in a loss of specificity. The CVs were empirically determined for each assay based on the information collected in the protocol optimization phases (Phase A & B) and the prevalidation. Thus, they can differ between assays.
In addition, we attach the document *Analytical procedure to identify and eliminate outlying observations* written by the responsible statistician, Sebastian Hoffmann, during the validation study and which gives rationales for applying this procedure.

5. Selection of test substances
e-mail of David Allen 12/01/2007 – question 3

Please find attached the file "Rationale for selection of test substances".

6. Removal of DMSO
e-mail of David Allen 12/01/2007 – question 4

Schindler et al 2004 state:

We sought to develop a protocol which would allow the use of the thawed whole blood samples directly without any washing steps to remove the cryoprotectant, as such a step would eliminate the essential advantages of the human whole blood assay, i.e., the ease of performance which allows a high degree of standardization as shown for various applications (Fennrich et al., 1999). Furthermore, besides stress and handling artifacts, the cells would lose their autologous plasma that permits a number of physiological responses, e.g., the sensitive response to

lipopolysaccharides (endotoxin, LPS) via lipopolysaccharide binding protein (LBP; Schumann, 1992; Fenton and Golenbock, 1998).

Indeed DMSO is not removed and up to now artefacts attributed to the presence of DMSO were not observed. The presence of DMSO enhances the IL-1 production and leads to a delay in the release. The fact that no wash step is required reduces strongly variation and introduction of artefacts.

7. Possible cytotoxicity
e-mail of David Allen 12/01/2007 – question 5

The aspect of cytotoxicity is covered by interference testing. As stated already on various occasions, interference testing (what we called positive product control in the validation study) is a must before you can use the WB (and the other) assays. If a substance would interfere with the assay by being cytotoxic, the spike recovery would be below 50%.

8. Freeze-thaw step for CRYO WB/IL-1 (Konstanz method)
e-mail of David Allen 12/01/2007 – question 6

This question was already posed during the drafting of the ICCVAM peer review documents (Mail David Allen 1/08/2006 question 3 and attached document PyroProtocol31Jul06)

In our reply (sent on 8/09/2006 with attachment reply_PryoProtocol31Jul06), we stated on page 2: *The freezing thawing enhances the IL-1 release and makes the Konstanz method more robust and reliable. It is not needed for the PEI method since the IL-1 release levels are higher.*
In fact, it should read that the freezing thawing enhances the IL-1 yield since the IL-1 produced in but not released by the monocytes is also measured.
It has been shown by Boneberg and Hartung (2003) that 10fold higher concentrations of (pro-)IL-1ß are found when including intracellular cytokine by whole blood lysis:

Ref *Boneberg E. and Hartung T. Febrile temperatures attenuate IL-1β release by inhibiting proteolytic processing of the proform and influence Th1/Th2 balance favoring Th2 cytokines. J. Immunol. 2003, 171:664-8.* attached.

Appendix C4

Rationale for the Selection of the 10 Substances Tested in the Validation/Catch-Up Validation Study of *In Vitro* Assays for Pyrogen Testing

[This Page Intentionally Left Blank]

Rationale for selection of the 10 substances tested in the validation/catch-up validation study of in vitro assays for pyrogenicity testing

Selection committee:
Thomas Montag-Lessing (chair), Michael Jahnke, Ingeborg Aarberge, Sandra Coecke

The main points which led to the selection were stability of the spikes, relevance, availability and costs of the substances:

1. Stability of the spikes, coding, interference testing

Experiments to evaluate the stability of endotoxin spikes in the final products revealed that stability of low endotoxin concentration could not be guaranteed over the time period needed for the prevalidation/validation study. Therefore, endotoxin spikes in higher (stable) concentrations were produced, filled in separate vials and coded. The laboratories received the clean substance plus the coded spikes, the clean substance had to be used for interference testing and contaminated with the coded spikes for the actual tests.

2. Relevance

The absence of pyrogens is crucial for intravenously administered drugs, this is reflected in the rabbit pyrogen test where the test substance is injected into the ear vene.
Therefore, only substances intended for i.v. injection were selected. In addition, it could be evaluated whether the in vitro assays would be able to detect 0.5 IU/ml endotoxin, which corresponds to threshold inducing fever in rabbits.

3. Availability/feasibility

- Substances should be on the market, thus the final product in the original vials could be tested and the conditions under which a lab performing final lot release would work could be met, e.g.
    - avoid possible contamination with pyrogens during opening the vials, drawing the samples etc
    - performing interference testing (corresponds to positive product testing in the validation trial)

- One lyophilised product was included in order to check for potential failures (e.g. pyrogenic contamination during reconstitution of the drug)

- Substance not interfering with any of the assays in order to control the correctness of the spiking procedure. Therefore, 0.9% NaCl pyrogen-free solution was included (Drug A and B)

4. Costs

- Due to the restricted funds available, costs of the substances to be tested in the validation trial played a role, e.g. it was not possible to include a rather expensive blood product as coagulation Factor VIII (Haemate® was used in the prevalidation trial).

Appendix C5

Comparison and Validation of Novel Pyrogen Tests Based on the Human Fever Reaction: Trial Data Report

[This Page Intentionally Left Blank]

Validation of Biomedical Testing Methods

Comparison And Validation Of Novel Pyrogen Tests Based On The Human Fever Reaction

Acronym: Human (e) Pyrogen Test

Trial data report

Date of Circulation:
File Identifier: *Stp-HPTVv03*

JOINT RESEARCH CENTRE
EUROPEAN COMMISSION
Institute for Health and Consumer Protection
ECVAM

TABLE OF CONTENTS Page no.

1 INTRODUCTION....3
2 THE BASIC BIOLOGICAL PRINCIPLE....4
3 STATISTICAL PROPERTIES OF THE DATA....5
4 METHODS....6
4.1 Background....6
4.2 Method A: Identification and elimination of aberrant data....6
4.3 Method B: Assuring the limit of detection....7
4.4 Method C: The prediction model....7
4.5 Method D: 2x2 contingency tables for the final results....9
5 ANALYTICAL PROCEDURE....10
5.1 General procedure....10
5.2 Modification for the test systems PBMC and WBT-NI....12
6 ADDITIONAL ANALYSIS....13
7 REFERENCES....13

Comparison And Validation Of Novel Pyrogen Tests Based On The Human Fever Reaction

Acronym: Human (e) Pyrogen Test

1 INTRODUCTION

The objective of the "Human(e) Pyrogen Test" project is to assess the performance and use of six recently developed *in vitro* pyrogen tests. These tests are based on the human fever reaction. As they are meant to be similar to the currently used Rabbit Test, the analytical procedure is designed to give a dichotomous outcome. In detail, drugs, which have to be tested for pyrogenic contamination due to regulatory instruction, have to be classified either as hazardous, i.e. pyrogenic, or as safe for humans. Hence, securing the safety of humans is the primary objective of pyrogenicity testing in general. Therefore, the prediction model is constructed to give a clear-cut classification of a given drug taking the safety aspect into account.

In addition to the prediction model, procedures to ensure quality criteria the test systems have to meet are included. A two-step method to identify and eliminate aberrant data as well as a test for a sufficient limit of detection are provided. Information from previous phases of the project established the basis to develop and define these methods. The data from the pre-validation study were used to refine the procedures.

All methods of the analytical procedure were developed to be applicable to each of the six test systems and were accepted by the participants.

.

2 THE BASIC BIOLOGICAL PRINCIPLE

The six test systems make use of the same biological principle. The mediators of the human fever reaction are cytokines, which are produced by monocytes. This principle is employed by incubating either fresh human cells or cell lines with the drug to be analysed under SOP-defined conditions. As there are several cytokines, which highly correlate with the human fever reaction, the cytokine of choice of the test sytems differs. Mainly the cytokine IL-6, but also IL-1β, TNF and neopterin were chosen as endpoints. After the incubation, an also SOP-defined ELISA-step is performed. In this step the cytokine is bound, visualised and finally measured by an optical reader. As the visualisation of the endpoint, measured as optical density (OD), is proportional to the amount of cytokine present, the resulting data are metrically scaled. In these entire procedure, a biological standard, WHO-LPS 94/580, is employed as an objective tool for comparison.

One of the test systems is based on a competitive ELISA, which results in a monotonically decreasing dose-response relationship, whereas the other systems show an increasing relationship due to their sandwich ELISA technique.

3 STATISTICAL PROPERTIES OF THE DATA

In previous phases the statistical properties of the data were analysed. Replicate observation for a fixed known control or an unknown drug revealed a right-skewed distribution. In experiments with large sample sizes it was shown, that a ln-transformation of the raw OD-data allows to assume a gaussian distribution of the data, which parameters can be estimated by the mean and the empirical variance.

As handling errors in the conduction of the test result in extreme observations, which may have an crucial impact on the prediction model, the probability of occurrence and impact of these observation was analysed. Although the probability of extreme observations is small for all tests, the inclusion of a method to identify and eliminate these data is indicated to ensure an optimised performance of the prediction model.

Furthermore, it was confirmed, that the dose-response relation ship between concentration of the contamination and the response increases, respectively decreases, monotonically for increasing concentration.

4 METHODS

4.1 Background

The entire analytical procedure consists of three different techniques, two of which assure the appropriateness of the data. The ELISA-plates employed have a 96-well format. The data of one plate have to be considered as a whole, which can not be compared to other plates due to uncontrollable variation. Therefore, each plate has to include all controls required for the analytical procedure. These are a negative control, which is 0.9%-NaCl, and a positive control of the WHO-LPS 94/580 standard diluted in 0.9%-NaCl, as well as negative and positive controls of the drugs, which are to be tested on the plate. Negative controls of a drug are obtained by released batches of the drugs. Positive controls are gained by adding 0.5 endotoxin units (EU)/ml of WHO-LPS 94/580. These 0.5 EU/ml were concordantly defined as the threshold concentration of endotoxin that induces fever in humans under worst conditions. In previous experiments it was shown, that this positive control lies in the most sensitive region, i.e. the steepest part, of the dose-response curve of all six test systems. In the following the NaCl-controls are denoted as "C-" (negative) and "C+" (positive). Similarly, the controls of a drug S_i are denoted as "S_i-" ans "S_i+". Furthermore, "S_ij", j=1, 2, 3, 4, 5, represent the blinded versions of the drug S_i.

4.2 Method A: Identification and elimination of aberrant data

The first method to be applied is an method to check the quality of the data of a plate. In general, this is done by a two-step procedure, which firstly identifies the sets of replicates with an extremely large variation. A set of replicates consists of four replicates per control, respectively drug tested. For every test system a maximal coefficient of variation (CV_{max}) was extracted from the available information. If the CV of a set of replicates is smaller than its CV_{max}, it is analysed as it is. Otherwise, the set is examined in the second step. This second step is a test for outliers. Therefore, the Dixon's test (1), which is USP approved, was chosen with the significance level of α=0.01. Preliminary to the testing itself, the raw OD-data are transformed with the natural logarithm, which normalises the data to meet the

prerequisites of the Dixon's test. If one observation in a set, which is identified by the Dixon's test, is responsible for its large variation, then this observation is excluded. If the variation is due to all observations, i.e. the absence of an outlier, the entire set of replicates is excluded from further analysis. Unfortunately, this approach poses the danger, that a whole plate can not be analysed, when a control is to be excluded. Therefore, both steps were chosen conservatively. Nevertheless, the empirical nature of the first step is not optimal and depends on general properties of the test system. But all established statistical methods, which address this problem, e.g. the Bartlett test for heterogeneity of variances, are not appropriate, because the variance structure over the range of concentration is highly variable and their global character. In table 1 the empirically derived CV_{max} are listed for the six test systems. The approach could be harmonized over all test systems.

test system	MM6	PBMC	THP-1BN	THP-1IK	WBT-KN	WBT-NI
CV_{max}	0.25	0.45	0.45	0.25	0.45	0.45

Table 1: Maximum CV's for the six test systems

4.3 Method B: Assuring the limit of detection

The second method is designed to ensure an minimum limit of detection of a plate (2). Because of the pre-defined dichotomous classification, a crude criterion, which merely shows strict monotonicity in the interesting part of the dose-response curve, can be chosen. Therefore, a one-sided t-test with a significance level of α=0.01 is applied to the ln-transformed data to ensure, that the response to the positive control is significantly larger than that of the respective negative control.

4.4 Method C: The prediction model

The third and most important statistical tool is the so-called prediction model (PM). In general, it is a statistical model, which classifies a given drug by an objective diagnostic or deciding rule. The objective of a dichotomous result requires a clear cut PM, which assigns a drug in one of the two classes "pyrogenic for humans" and "non-pyrogenic for humans". As the members of the project decided on a threshold positive

control, a one-sided test is appropriate for the task. Because the data are normalised by a ln-transformation, a t-test was chosen. Although the variances over the range of concentration converge by the transformation, the assumptions of equal variances does generally not hold true, because it depends on additional covariates. Therefore, the one sided Welch-t-test (3) is applied. Due to the safety aspect of the basic problem, the hypotheses of the test are

$$H_0 : \mu_{S_i j} > \mu_{S+} \quad vs \quad H_1 : \mu_{S_i j} < \mu_{S+},$$

where $\mu_{...}$ denotes the parameter of location of the respective ln-transformed distribution. This approach controls the probability of false positive outcomes directly by means of its significance level α, which is chosen as 0.01, because is assumes hazard, respectively pyrogenicity, of the tested drug in H_0, and assures safety, i.e. non-pyrogenicity. The test statistic is

$$T_{S_i j} = \frac{\bar{x}_{S+} - \bar{x}_{S_i j}}{\sqrt{\frac{s^2_{S+}}{n_{S+}} + \frac{s^2_{S_i j}}{n_{S_i j}}}}.$$

The PM is built by means of the outcome of the test. Let 0 denote safety and 1 denote hazard. The classification of S_i-j is then determined by

$S_i j = 0$, if $T_{S_i j} > t_{0.99; n_{S+} + n_{S_i j} - 2}$,

$S_i j = 1$, else,

where $t_{0.99; n_{S+} + n_{S_i j} - 2}$ the 0.99-quantile of the t-distribution with $n_{S+} + n_{S_i j} - 2$ degrees of freedom. The number of replicates for every control and sample, i.e. $n_{...}$, was harmonised for all test systems to be four. Due to the possibility of removing one observation by the outlier test, the number of replicates could be reduced to three. The classification of a version of a drug is regarded as an independent decision. Therefore, the niveau α is local.

4.5 Method D: 2x2 contingency tables for the final results

Finally, the classifications of the drugs will be summarised in 2x2 contingency tables, formally presented in table 2.

		pre-defined class		Σ
		1	**0**	
classification by test system and PM	**1**	**a**	**b**	**a+b = $n_{.1}$**
	0	**c**	**d**	**c+d = $n_{.0}$**
Σ		**a+c = $n_{1.}$**	**b+d = $n_{0.}$**	**n**

Table 2: 2x2 contingency table

From these tables estimates of the sensitivity, i.e. the probability of correctly classified positive drugs, and specificity, i.e. the probability of correctly classified negative drugs, will be obtained by the respective proportions. Furthermore, these estimates will be accompanied by confidence intervals, which will be calculated by the Pearson-Clopper method (4). For example, let $\hat{p}_{SE}$ denote the proportion, namely the sensitivity, under investigation. Then the confidence interval to a niveau α is calculated as

$$\left[p_{SE}^{L} = \frac{\mathrm{a} F_{2\mathrm{a};2(n_{1.}-\mathrm{a}+1);\frac{\alpha}{2}}}{n_{1.}-\mathrm{a}+1+\mathrm{a}F_{2\mathrm{a};2(n_{1.}-\mathrm{a}+1);\frac{\alpha}{2}}}; p_{SE}^{U} = \frac{(\mathrm{a}+1)F_{2(\mathrm{a}+1);2(n_{1.}-\mathrm{a});1-\frac{\alpha}{2}}}{n_{1.}-\mathrm{a}+(\mathrm{a}+1)F_{2(\mathrm{a}+1);2(n_{1.}-\mathrm{a});1-\frac{\alpha}{2}}} \right],$$

where $F_{\ldots}$ denotes the respective quantile of the F-distribution and $n_{1.}$ is the sample size of the positive drugs and a the number of correctly classified drugs. By contaminating the drugs artificially and by defining a threshold dose, which is assumed to be appropriate, the class of a drug is determined beforehand. The versions of drugs, which are effectively contaminated, but below the threshold dose, are considered to be negative, respectively safe, because their contamination is not crucial for humans in terms of endotoxin limit concentration.

5 ANALYTICAL PROCEDURE

5.1 General procedure

The process of the analytical procedure is highlighted in figure 1. Firstly, the data of the controls of a ELISA-plate are checked for aberrant data with procedure A. If indicated, outliers are removed. If sets of replicates are to be removed, this is recorded, but due to the empirical base of the first step of A, the data will be further analysed with reservations. Afterwards, the remaining data of the controls are tested with method B to ensure a minimum limit of detection. If the controls in 0.9%-NaCl do not differ significantly, the further analysis is done with reservations. If the controls of a drug do not differ significantly, all data of this drug do not qualify for further analysis. The last part of quality assurance is the application of method A to the data of the blinded drugs. Here, drugs, which fail the criteria, are removed from further analysis. Finally, the remaining data are put to the prediction model. The classification of the still blinded drugs are sent to ECVAM in an official document, which will in general comprise the assigned class for every drug structured by test system, laboratory and drug. Upon receipt ECVAM will send the blinding code in a electronically generated document by e-mail to the project's statistician Sebastian Hoffmann (e-mail: sebastian.hoffmann@uni-konstanz.de). Additionally, a hardcopy of the blinding code will be sent by post. Once the data are unblinded, the final results, which core will be method D, can be summarised, explicitly analysed and appropriately presented. Additionally to the contingency tables and related topics, an inter-laboratory comparison will be done. Furthermore, the reasons for misclassifications will be identified.

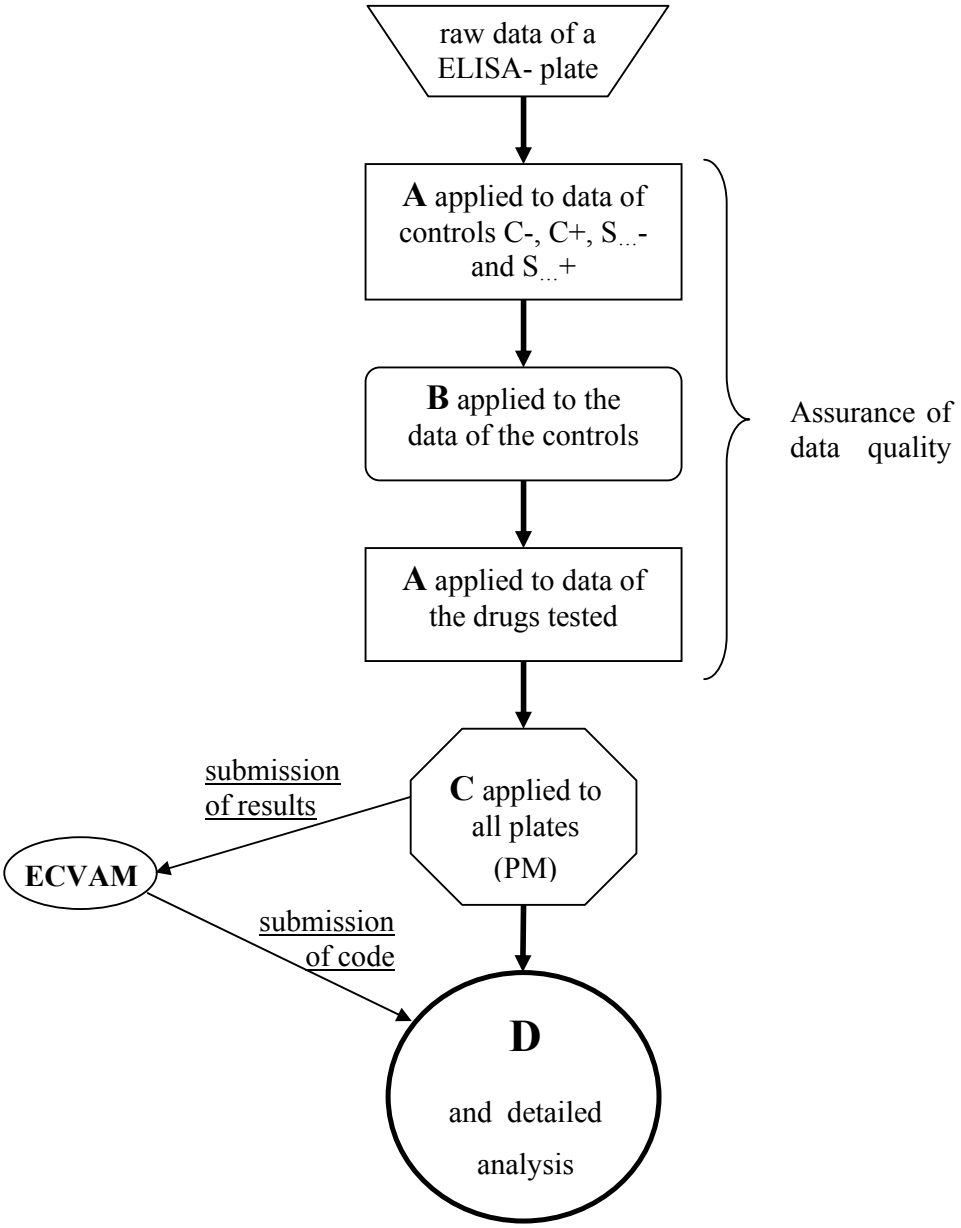


Figure 1: Flowchart of the main analytical procedure

5.2 Modification for the test systems PBMC and WBT-NI

This analytical procedure has to be modified for the two test systems PBMC and WBT-NI. The two test systems base their classification of a drug not on one outcome, but they classify a drug by the results of several independent experiments. Because both methods rely on fresh blood, the inter-donor variability is taken into account by using the blood of several donors and conducting the test independently. The modifications, which have to be made, arise out of contradictory classification of a drug by different donors. Therefore the classification of a drug is determined by the combination of the single donor-dependent results, which are calculated with the analytical procedure presented.

6 ADDITIONAL ANALYSIS

The detailed analysis will contain an inter-lab comparison per test system, whereas measures of correlation as well as similarity can be used. Furthermore, shortcomings depending on specific drugs, laboratories, the analytical procedure and/or test systems will be examined by exploratory statistical methods.

Additionally, modifications of the methods A, B and C will be used to optimise the analytical procedure with the information from the new data. In general, these are the consequences of more restrictive or less restrictive assumptions. With regard to A, the results of a procedure without a tool for aberrant data will be compared to the results of the described procedure. For B, more restrictive criteria to ensure a valid dose-response relationship will be applied, e.g. techniques for ratios between controls based on Fieller's theorem (5, 6). Besides, modifications in the t-test of the prediction model will be of interest, mainly assumptions considering the variance and a multiple testing approach. E.g. a simulation, which allows for the $\sqrt{k}$-rule optimising the Dunnett's test could be realisable.

Finally, methods taking the real life situation of pyrogen testing into account will be highlighted. These include a Fieller-based method to handle interference.

7 REFERENCES

1. Barnett, V., Lewis, T. (1984). *Outliers in statistical data*. 2nd edition, Chichester, US: John Wiley & Sons, pp 171-172.
2. Hothorn, L. A. (1995). Biostatistical analysis of the 'control vs. k treatments' design including a positive control group. In *Testing principls in clinical and preclinical trials*, pp 19-26, Stuttgart, Germany: Gustav Fischer Verlag.
3. Snedecor, G. W., Cochran, W. G. (1989). *Statistical methods*, 8th edition, Ames, US: The Iowa State University Press.
4. Clopper, C. J. & Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26**, 404-413
5. Fieller E. (1954) Some problems in interval estimation. *J. Royal Statistical Society* **B16**, 175-185.
6. Hauschke, D., Hothorn, T., Schaefer, J. (2002). The role of control groups in mutagenicity studies: matching biological and statistical relevance. *ATLA* (in press)

Appendix C6

List of Drugs for the Catch-Up Validation Study

[This Page Intentionally Left Blank]

VALIDATION STUDY (CATCH UP): LIST OF DRUGS

Product	Manufacturer	Lot
Alkohol-Konzentrat 95%[1]	B. Braun	2465Z01
Beloc i.v.	Astra Zeneca	DA419A1
Binotal 0,5g	Grünenthal	117EL2
Fenistil	Novartis	26803[2]
MCP Hexal	Hexal	21JX22
Orasthin[3]	Hoechst	not available
Sostril	Glaxo Wellcome	3H01N[4]
Syntocinon 3 I.E.	Novartis	S0040
Traubenzuckerlösung 5% Eifelfango[5]	Eifelfango	3132[6]

16.02.2004

[1] "95% Alcohol Concentration"
[2] Different lot number
[3] Orasthin no longer on the market, replaced with Syntocin 3 I.E. containing also oxytocin
[4] Different lot number
[5] "5% Glucose Solution"
[6] Different lot number

[This Page Intentionally Left Blank]

Appendix C7

Analytical Procedure to Identify and Eliminate Outlying Observations

[This Page Intentionally Left Blank]

Analytical procedure to identify and eliminate outlying observations

Introduction

As seen in the pre-validation, the problem of outlying observations is not appropriately solved yet. The crucial issue about these observations is their impact on the prediction model, which could result in false classifications of substances. Of course, one could just neglect such observation, as done in the pre-validation. This is the most easy way, but its appropriateness with regard to sensitivity and specificity is depending on the probability of outlying observations. So far, this probability was low, smaller than 5%, for all assays. Nevertheless, even if there are only a few outliers, this approach has the disadvantage, that one would have to live the most extreme and maybe even pre-identified outlying observations, e.g. when a technician recognises that she/he made a gross handling error.

Therefore, a new analytical procedure was developed. First of all, the objective of such an procedure has to be defined precisely. On the one hand, a way to identify obvious handling errors, which most often can be identified by eye by trained persons, is needed. On the other hand, a method to handle sets of replicates, which are extremely untypical for a specific assay, has to be taken into account. Hence, a generally applicable two-step procedure is proposed.

Step 1: Checking the variation between sets of replicates

Firstly, the data of an ELISA-plate are checked for untypical variation of one or more sets of replicates. In the given situation, one is only interested in those sets with extremely large variance. In general, there are two situations which have to be considered. On the one hand, just one observation could be responsible for a huge variance in its set of replicates. On the other hand, equally distributed replicates over a large range of response, which includes the situation of two outliers in a set, might be the reason.

The existing statistical tests addressing this question, e.g. the Bartlett-test, are not appropriate due to various reasons, but mainly because they assume homogeneity of variances and show global heterogeneity. Thus a simple empirical method was derived for every assay, which is mainly based on the data of the pre-validation and the information from Phases A and B. The core of this method is an appropriate measure of variation for a set of replicates. Here, the CV is chosen, but the variance or the standard deviation can be used more or less equivalently. From these empirical information, a maximum CV, denoted as CV_{max}, was derived, which can be used as a tool to assess the variation of each set of replicates very easily. If a CV of a given set is larger than CV_{max}, then this set will be examined further in the second step of the

procedure. If the CV is smaller, then the data of the set will be analysed as they are. To highlight this empirical method, it is exemplarily explained for the Novartis-PBMC assay. In figure 1 the variation within sets of replicates for all available data is presented.

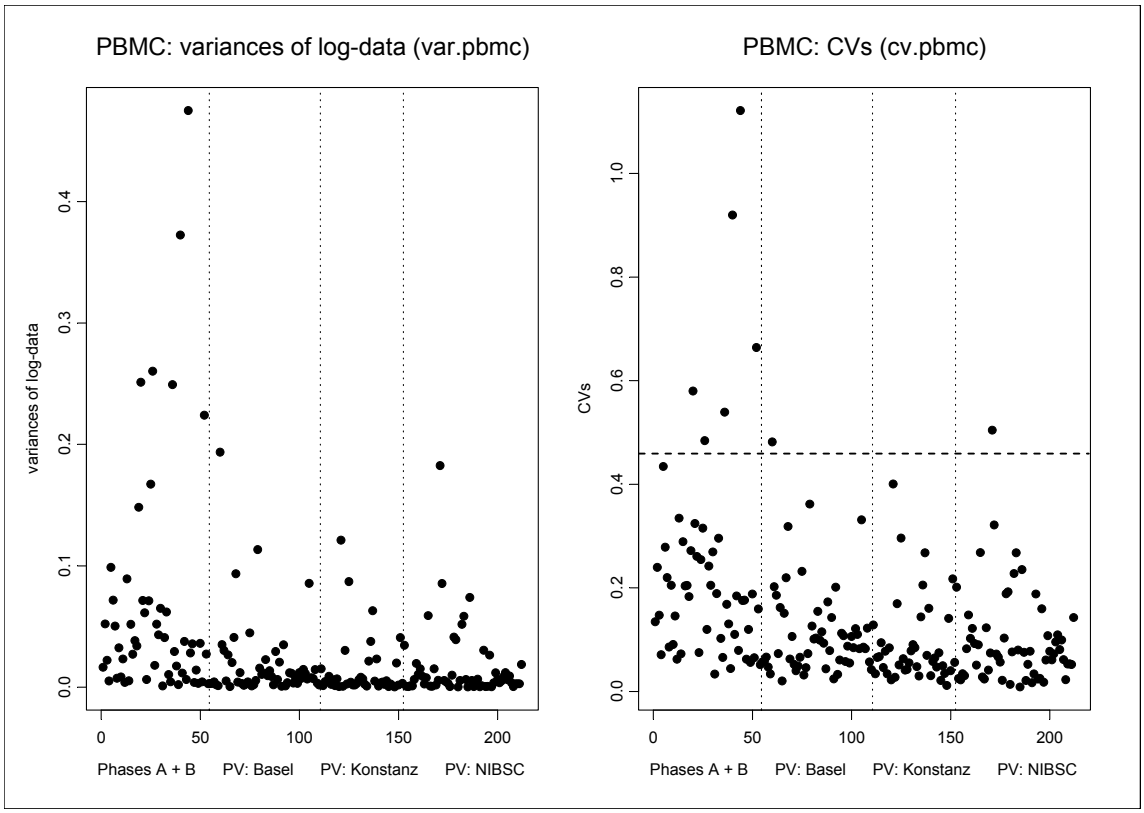


Figure 1: Variation within replicates for the PBMC-assay

Together with the raw data and some linear modelling techniques, here $CV_{max}(PBMC) = 0.45$ was chosen. This choice identifies two out of 154 pre-validation data sets as outlying sets in the right part of figure 1. As can be seen in the left part of figure 1, a criterion based on the variance, e.g. $Var_{max}(PBMC) = 0.18$ is almost identical, which even can be shown by some statistical approximation under certain assumptions.

Additionally to the approach with the CV, a criterion based on the ratio of variances was applied. Also having the empirical background, it did not show any advantage.

Step 2: Checking the variation within sets of replicates identified in step 1

Let S_i, i=1,…, denote the crucial, in step 1 identified sets with $CV(i) > CV_{max}$. In this second step the reasons for the high variation of the S_i's are examined. Firstly, a common test for outliers, the FDA-approved Dixon-test, is applied to each S_i with the niveau α of 1%. If an outlier is identified, it is withdrawn from its set and the remaining data are further analysed. If

no outlier is detected, the observations of a set of replicates are regularly distributed over a large range of response. In the latter case, it is recommended, to repeat the substance(s) S_i on another plate.

Discussion

Assuming such a partly empirical approach is appropriate, one still has to be aware of its properties and effects, especially when applied in the validation study. Firstly, the procedure gives excellent results when applied to the pre-validation data. This is expected, because the CV_{max} criterion was mainly derived by the data themselves, which makes it a self-fulfilling prophecy. Therefore, it poses the danger of choosing the CV_{max} too small, because it may lead to a lot of rejected sets. Additionally, maintaining such an empirical procedure demands to check regularly for the validity of the chosen $CV_{max.}$

Secondly, in the given situation of the validation study, the impact of the retrospectively applied procedure has to be taken into account. Because the data are checked outside the labs, the sample size of the number of classified samples could be reduced during analysis. For example, assume that a control on a plate, on which the prediction model is based (e.g. the positive product control), does fail the above proposed procedure. In the case of the positive product control, this would mean that none of the samples tested on that plate could be classified by the prediction model.

Furthermore, the robustness of the procedure with regard to systemic errors is noteworthy. It will work, even if the ELISA-plate is of low quality, e.g. with regard to coating, or if moderate systemic handling errors are present.

Application of the procedure to the available data

assay	CV_{max}	number of sets	Outlier procedure		Dixon test
			outliers	outlying sets	
THP-Bern	0.45	138	1	4	7
THP-Inns.	0.2	112	-	-	3
MM6	0.2	129	1	1	5
PBMC	0.45	154	1	1	6
WBT-Konst.	0.45	138	-	1	3

Table 1: Results of the outlier procedure

The results in table 1 are very promising, but should not be overestimated as mentioned above. For example, the problematic first run from Oslo with the THP-Bern assay with regard to Haemate can easily be handled with the proposed procedure, because three of the Haemates would have to be retested. In contrast, the Dixon test alone would not have detected any outlier in the three Haemates.

As can be seen in table 1, a harmonised choice for CV_{max} was sought. Alternatively, the more conservative $CV_{max} = 0.25$ for the two-plate cell line assays (THP-Innsbruck, MM6) could have been applied giving very similar results. The more restrictive $CV_{max} = 0.4$ for the THP-Bern and the two methods based on fresh blood could have also been chosen.

Unfortunately, the variation within sets of replicates for the WBT-NIBSC increased from Phases A and B to the pre-validation and is fortunately decreasing at the moment due to changes in the SOP. But considering the variation shown in Phases A and B and the harmonising aspect of the above proposal, a CV_{max}(WBT-NIBSC) = 0.45 seems to be appropriate.