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 This BRD includes data from previously conducted validation studies and from previously published and unpublished data.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This ICCVAM BRD also summarizes information from published studies and additional unpublished data provided by ECVAM.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.