A UNITED STATES DEPARTMENT OF COMMERCE PUBLICATION

NBS SPECIAL PUBLICATION 260-36

Standard Reference Materials:

A REFEREE METHOD FOR THE DETERMINATION OF CALCIUM IN SERUM

U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards

Standard Reference Materials:

A Referee Method for the Determination of Calcium in Serum

J. Paul Cali, John Mandel, and

Larry Moore

Institute for Materials Research
National Bureau of Standards
Washington, D.C. 20234

and

Donald S. Young

National Institutes of Health
Department of Health, Education and Welfare
Washington, D.C. 20214

U.S. DEPARTMENT OF COMMERCE, Peter G. Peterson, *Secretary*

Issued May 1972

Library of Congress Catalog Card Number: 72-600082

National Bureau of Standards Special Publication 260-36

Nat. Bur. Stand. (U.S.), Spec. Publ. 260-36, 136 pages (May 1972)

CODEN: XNBSAV

Issued May 1972

Foreword

The "260 Series" of NBS Special Publications was originally initiated and reserved for the purpose of reporting technical work done at NBS in the area of Standard Reference Materials. The scope of this program was recently enlarged to include support in the development of referee methods of analysis, i.e., methods of known accuracy based in part on the use of SRM's. It is now recognized that a meaningful measurement system comes into being only when its several parts are brought into a harmonious and relevant relationship. The development of referee methods of analysis is a necessary step in this process. These matters are discussed in detail in Sections I and II in the body of this report.

In Clinical Chemistry, referee methods, as defined herein, are lacking. This work is a first attempt to bring about a referee method of analysis in this field. A careful study of the methodology actually used will reveal some shortcomings, but it may be said with some assurance, subject to the many limitations and provisos listed, that a referee method for the analysis of calcium in serum or physiological solutions has been achieved. It may be argued that additional work to establish the magnitude of the several individual systematic biases could have been carried out, or that additional exercises should have been run, etc. However, when work of this nature is dependent to a large degree upon the voluntary participation of the cooperating laboratories and the free time and services of the Experts Committee then there must be a reasonable limit as to the total duration of the project. The work reported covered the period from June 1970 to December 1971. Furthermore, the Experts Committee felt that a reasonable goal had in fact been achieved, based on the work to date, and that further additional studies would not result in any marked improvement in the accuracy of the method.

Several of the findings are subjectively based and cannot be rigorously or fully supported by the cold evidence, e.g., the requirement that "well-motivated personnel, traine in the proper techniques are essential if the accuracy inherent in the method is to be achieved." We have no hesitancy in making such statements other than they are likely to produce some rather heated controversy, because this work is to be considered, in no small measure, as an educational exercise. Automation in clinical chemistry, or any other analytical field, cannot substitute for the technical skills and knowledge necessary to the practice of analytical chemistry--and clinical chemistry, in practice, is clinical analytical chemistry.

A word of caution is necessary. Referee methods have been called, at various times and for various purposes, standard methods, standard procedures, reference methods, validation methods, and a variety of other names. Because the principal purpose is to achieve accuracy in measurement, the term, referee method, is preferred over any other, to avoid implications of a legally imposed method, a preferred method, a method of qualification and the like. Although it is true that a referee method might, indeed, serve any one or more of these purposes, the firm intent of all concerned in developing the methodology described herein is wholly scientific.

Finally, it should be understood that the "260 Series" permits reporting in much more detail than is usually allowable in scientific journals. Again, the reason for this is largely educational. Very often in condensed versions a great deal of expertise, experience, and technique is assumed to reside in or be available to the reader. In this report we make a minimum of such assumptions. Furthermore a degree of editorial comment and expounding of the philosophical basis of the work is encouraged, and we have taken advantage of this principle to include an introduction

on the basics of meaningful measurement and the methodology required to develop referee methods.

PREFACE

Standard Reference Materials (SRM's) as defined by the National Bureau of Standards are "well-characterized materials, produced in quantity, that calibrate a measurement system to assure compatability of measurement in the nation." SRM's are widely used as primary standards in many diverse fields in science, industry, and technology, both within the United States and throughout the world. In many industries traceability of their quality control process to the national measurement system is carried out through the mechanism and use of SRM's. For many of the nation's scientists and technologists it is therefore of more than passing interest to know the details of the measurements made at NBS in arriving at the certified values of the SRM's produced. An NBS series of papers, of which this publication is a member, called the NBS Special Publication - 260 Series is reserved for this purpose.

This 260 Series is dedicated to the dissemination of information on all phases of the preparation, measurement, and certification of NBS-SRM's. In general, much more detail will be found in these papers than is generally allowed, or desirable, in scientific journal articles. This enables the user to assess the validity and accuracy of the measurement processes employed, to judge the statistical analysis, and to learn details of techniques and methods utilized for work entailing the greatest care and accuracy. It is also hoped that these papers will provide sufficient additional information not found on the certificate so that new applications in diverse fields not foreseen at the time the SRM was originally issued will be sought and found.

Inquiries concerning the technical content of this paper should be directed to the author(s). Other questions concerned with the availability, delivery, price, and so forth will receive prompt attention from:

Office of Standard Reference Materials
National Bureau of Standards
Washington, D.C. 20234

J. Paul Cali, Chief
Office of Standard Reference Materials

OTHER NBS PUBLICATIONS IN THIS SERIES

NBS Spec. Publ. 260, Catalog of Standard Reference Materials, July 1970. 75 cents.* (Supersedes NBS Misc. Publ. 260, January 1968 and NBS Misc. Publ. 241, March 1962.)

NBS Misc. Publ. 260-1, Standard Reference Materials: Preparation of NBS White Cast Iron Spectrochemical Standards, June 1964. 30 cents.*

NBS Misc. Publ. 260-2, Standard Reference Materials: Preparation of NBS Copper-Base Spectrochemical Standards, October 1964. 35 cents.*

NBS Misc. Publ. 260-3, Standard Reference Materials: Metallographic Characterization of an NBS Spectrometric Low-Alloy Steel Standard, October 1964. 20 cents.* (Out of print).

NBS Misc. Publ. 260-4, Standard Reference Materials: Sources of Information on Standard Reference Materials, February 1965. 20 cents.* (Out of print).

NBS Misc. Publ. 260-5, Standard Reference Materials: Accuracy of Solution X-Ray Spectrometric Analysis of Copper-Base Alloys, March 1965. 25 cents.* (Out of print).

NBS Misc. Publ. 260-6, Standard Reference Materials: Methods for the Chemical Analysis of White Cast Iron Standards, July 1965. 45 cents.*

NBS Misc. Publ. 260-7, Standard Reference Materials: Methods for the Chemical Analysis of NBS Copper-Base Spectrochemical Standards, October 1965. 60 cents.*

NBS Misc. Publ. 260-8, Standard Reference Materials: Analysis of Uranium Concentrates at the National Bureau of Standards, December 1965. 60 cents.* (Out of print).

NBS Misc. Publ. 260-9, Standard Reference Materials: Half Lives of Materials Used in the Preparation of Standard Reference Materials of Nineteen Radioactive Nuclides Issued by the National Bureau of Standards, November 1965. 15 cents.*

NBS Misc. Publ. 260-10, Standard Reference Materials: Homogeneity Characterization on NBS Spectrometric Standards II: Cartridge Brass and Low-Alloy Steel, December 1965. 30 cents.*

NBS Misc. Publ. 260-11, Standard Reference Materials: Viscosity of a Standard Lead-Silica Glass, November 1966. 25 cents.*

NBS Misc. Publ. 260-12, Standard Reference Materials: Homogeneity Characterization of NBS Spectrometric Standards III: White Cast Iron and Stainless Steel Powder Compact, September 1966. 20 cents.*

NBS Misc. Publ. 260-13, Standard Reference Materials: Mössbauer Spectroscopy Standard for the Chemical Shift of Iron Compounds, July 1967. 40 cents.*

NBS Misc. Publ. 260-14, Standard Reference Materials: Determination of Oxygen in Ferrous Materials -- SRM 1090, 1091, and 1092, September 1966. 30 cents.*

NBS Misc. Publ. 260-15, Standard Reference Materials: Recommended Method of Use of Standard Light-Sensitive Paper for Calibrating Carbon Arcs Used in Testing Textiles for Colorfastness to Light, June 1967. 20 cents.*

NBS Spec. Publ. 260-16, Standard Reference Materials: Homogeneity Characterization of NBS Spectrometric Standards IV: Preparation and Microprobe Characterization of W-20% Mo Alloy Fabricated by Powder Metallurgical Methods, January 1969. 35 cents.*

NBS Spec. Publ. 260-17, Standard Reference Materials: Boric Acid; Isotopic and Assay Standard Reference Materials, February 1970. 65 cents.*

NBS Spec. Publ. 260-18, Standard Reference Materials: Calibration of NBS Secondary Standard Magnetic Tape (Computer Amplitude Reference) Using the Reference Tape Amplitude Measurement "Process A", November 1969. 50 cents.*

NBS Spec. Publ. 260-19, Standard Reference Materials: Analysis of Interlaboratory Measurements on the Vapor Pressure of Gold (Certification of Standard Reference Material 745), January 1970. 30 cents.*

NBS Spec. Publ. 260-20, Standard Reference Materials: Preparation and Analysis of Trace Element Glass Standards. (In preparation)

NBS Spec. Publ. 260-21, Standard Reference Materials: Analysis of Interlaboratory Measurements on the Vapor Pressures of Cadmium and Silver, January 1971. 35 cents.*

NBS Spec. Publ. 260-22, Standard Reference Materials: Homogeneity Characterization of Fe-3Si Alloy, February 1971. 35 cents.*

NBS Spec. Publ. 260-23, Standard Reference Materials: Viscosity of a Standard Borosilicate Glass, December 1970. 25 cents.*

NBS Spec. Publ. 260-24, Standard Reference Materials: Comparison of Redox Standards, January 1972. $1.00.*

NBS Spec. Publ. 260-25, Standard Reference Materials: A Standard Reference Material Containing Nominally Four Percent Austenite, February 1971. 30 cents.*

NBS Spec. Publ. 260-26, Standard Reference Materials: National Bureau of Standards-U.S. Steel Corporation Joint Program for Determining Oxygen and Nitrogen in Steel, February 1971. 50 cents.*

NBS Spec. Publ. 260-27, Standard Reference Materials: Uranium Isotopic Standard Reference Materials, April 1971. $1.25.*

NBS Spec. Publ. 260-28, Standard Reference Materials: Preparation and Evaluation of SRM's 481 and 482 Gold-Silver and Gold-Copper Alloys for Microanalysis, August 1971. $1.00.*

NBS Spec. Publ. 260-29, Standard Reference Materials: Calibration of NBS Secondary Standard Magnetic Tape (Computer Amplitude Reference) Using the Reference Tape Amplitude Measurement "Process A-Model 2", June 1971. 60 cents.*

NBS Spec. Publ. 260-30, Standard Reference Materials: Standard Samples Issued in the USSR (A Translation from the Russian), June 1971. $1.00.*

NBS Spec. Publ. 260-31, Standard Reference Materials: Thermal Conductivity of Electrolytic Iron SRM 734 from 4 to 300K, November 1971. 35 cents.*

NBS Spec. Publ. 260-32, Standard Reference Materials: The Cooperative Study of Temperature Scale Standards for DTA by ICTA and NBS. (In preparation)

NBS Spec. Publ. 260-33, Standard Reference Materials: Comparison of Original and Supplemental SRM 705, Narrow Molecular Weight Distribution Polystyrene, H. L. Wagner. (In preparation).

NBS Spec. Publ. 260-34, Standard Reference Materials: Thermoelectric Voltage. (In preparation).

NBS Spec. Publ. 260-35, Standard Reference Materials: Thermal Conductivity of Austenitic Stainless Steel, SRM 735 from 5 to 280K, April 1972. 35 cents.*

NBS Spec. Publ. 260-36, Standard Reference Materials: A Referee Method for the Determination of Calcium in Serum. SRM 736.(This publication)

NBS Spec. Publ. 260-37, Standard Reference Materials: Methods of Analysis of NBS Clay Standards. (In preparation).

NBS Spec. Publ. 260-38, Standard Reference Materials: Preparation and Calibration of Standards of Spectral Specular Reflectance. (In preparation).

NBS Spec. Publ. 260-39, Standard Reference Materials: The Eddy Current Decay Method for Resistivity Characterization of High-Purity Metals. (In preparation).

*Send order with remittance to: Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Remittance from foreign countries should include an additional one-fourth of the purchase price for postage.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

PAGE

Figure No.

A REFEREE METHOD FOR THE DETERMINATION OF CALCIUM IN SERUM

J. Paul Cali, John Mandel, and Larry Moore

Institute for Materials Research
National Bureau of Standards
Washington, D.C. 20234

and

Donald S. Young

National Institutes of Health
Department of Health, Education and Welfare
Washington, D.C. 20214

The first referee method of analysis (a method of known accuracy) in clinical chemistry has been developed by a team of experts from clinical chemistry, 8 cooperating clinical chemistry laboratories, and scientists from the National Bureau of Standards. Subject to the limitations and provisos stated, calcium in serum can be determined by the referee method to within ±2% of the "true or absolute" value. Also discussed are: (1) What Constitutes a Meaningful Measurement System; (2) Theory of Operation and Practice Required to Develop a Referee Method; (3) History of the Actual Work; the Referee Method Protocol; (4) Results; (5) The Isotope-Dilution Mass Spectrometry Work; (6) Discussion, Conclusions, and Recommendations for Future Referee Method Developments.

Key words: Accuracy; atomic absorption spectrometry; calcium in serum; clinical chemistry; referee method for calcium.

I. INTRODUCTION

A. A Meaningful Measurement System

In 1966, Vannevar Bush said, "If men are to accomplish together anything useful whatever, they must, above all, be able to understand one another." Understanding between men implies communication, and in science and technology, effective communication between and among scientists, engineers, technicians, and all those involved in the various aspects of scientific and technological endeavors, is realized, in large part, through the measurement process. However, the degree to which the communication is useful and effective depends largely on whether the measurements made are in actuality meaningful. When a measurement is meaningful, the value or number associated with the property under consideration meets three criteria--the number obtained during measurement truly reflects in some well-defined way the property being measured, and the number is both accurate and precise. In ideal cases, and certainly in "pure" measurement science, a direct relationship can be traced, usually quite straightforwardly between the property measured and the base units of the Systeme International. The realization of the candela, insofar as the material aspects of the measurement are concerned depends only on the purity of the platinum since the candela is defined in terms of that material. In this instance, meaningful measurement (or realization) of the candela depends upon knowing the purity of the platinum to be used to a high degree of accuracy and precision.

As we move down the ladder from the realm of "pure" measurement science into the murkiness of applied measurement, the task of making meaningful measurement becomes ever increasingly difficult. In the United States in 1971 more than 10^9 measurements were made in over 20,000 clinical laboratories. Yet if the three criteria mentioned above

are applied to these measurements, then it is certain that a large fraction, perhaps more than 25%, are not meaningful at all.

It is instructive to examine in some detail what constitutes a meaningful measurement system. Too often scientists and engineers concentrate on one aspect of a system, not realizing that unless all parts of the system are tended and cared for, that the hoped-for-result will not be achieved.

A universal, meaningful measurement system consists of five distinct parts:

1. A rational, self-consistent system of units of measurement. (E.g., the S.I., now universally accepted.
2. The means to realize in practice the defined units and their derivatives, for example, through the use of certified standard reference materials. In areas of applied technology the production and availability of well-characterized materials is of utmost importance in allowing the transfer of meaningful measurement by these means from highly competent laboratories that measure and produce these materials, to the user in the applied sector.
3. Methods of measurement or analysis or test, based upon the well-characterized materials of part 2, and which are, above all, accurate methods. Only in a system based on accuracy can communication between different fields of science and technology be effective. Precision alone is insufficient. These basic, accurate methods are often referred to as referee methods, reference methods, or standard methods. In clinical chemistry the first term, referee methods, is the phrase of choice. In any case, they should be methods of proven accuracy. They need not be, nor are they often in practice, the methods applied or used in industrial or technological application.

They, together with the well-characterized materials of part 2, are available to test and check the methods of part 4, below.

4. Field or applied methods of measurement or analysis or test. These are the methods applied on a large scale in everyday work. They have been introduced, often at the expense of accuracy, to turn out a large volume of results. However, if the methods of part 3 are available, then there is no fundamental reason why these practical methods should not be checked or validated against the referee method to assure their accuracy. Of course, if the referee method of part 3 can be made to meet the requirements of the field methods (speed, ease of operation, low-cost, etc.) then the referee method and field method should be made identical.

5. A method whereby the long-term integrity of the measurement system is assured. Given that parts 1-4 are operative and under control, then a mechanism is set up whereby the various practitioners and users of the system are periodically tested against materials supplied by a checking or validating agency (industrial, governmental, professional society, etc.). The properties of the material to be tested are known only to the checking agency via parts 2 and 3. Returned results are statistically analyzed and it is easily determined whether the measurement process in individual laboratories or an entire industry is in fact under control.

Obviously, the implementation and operation of a meaningful measurement system for all science, industry and technology is far beyond the ability and resources of any one organization. The work required is so vast in scope that only through a large network of qualified and inter-

ested organizations can order be brought to bear. In the steps outlined above, the organizations listed below should bear some responsibility for implementing and/or continuing their efforts to provide the required resources to bring about meaningful measurement:

Part 1 - International Committee of Weights and Measures (CIPM); National Standards Laboratories; International Union of Pure and Applied Chemistry (IUPAC).

Part 2 - National Standards Laboratories; Qualified Manufacturers; IUPAC and CIPM (information dissemination); Professional Societies.

Part 3 - Professional Societies; Standards Bodies (e.g., ISO, ASTM); Governmental Agencies and Laboratories.

Part 4 - Commercial Laboratories and Associations; Standards Bodies; Professional Societies.

Part 5 - Governmental Agencies and Laboratories; Commercial Laboratories and Associations; Standards Bodies.

B. Why Accuracy in Clinical Chemistry

In the preceding section, stress was laid on accuracy in measurement. Some confusion still exists as to the exact meaning of accuracy, especially as opposed to precision. One of the best discussions on this subject was given by C. Eisenhart [1]. A paper titled "The Attainment of Accuracy in Clinical Analytical Methods" applies more particularly to the field of clinical chemistry [2]. When the value obtained by the measurement process for the property under consideration is essentially the same as that obtained by any other acceptable measurement process, or, alternatively, is the value predicted by a well-developed theory, then accuracy has been achieved. The key word in the above is "acceptable". An acceptable measurement process is one the accuracy of which has been established by some other independent route, and its remaining systematic biases (never

totally reduced or even known with certainty) are small relative to the accuracy goal of the measurement process under test. Clearly, in practice, there will be many instances where neither an acceptable measurement process is available nor where theory can be relied on to assure accuracy. In such cases, consensus, by experts, is probably the only viable alternative. Such a route, however, should be considered a stop-gap measure and additional work performed to establish an independent method of known accuracy.

The question often arises among scientists in clinical chemistry, "Why is accuracy needed, in any case, as in practice, very large uncertainties in the overall, total medical picture remain?" There are several answers to this question. First, there is a legal argument. No matter how precise a method is shown to be, the values obtained by a precise but inaccurate method cannot and will not be sustained under legal challenge if values obtained by a method shown to be more accurate are offered in evidence. A more accurate value is by definition "closer to the truth," and would tend to carry more weight in legal disputes.

Second, and more important from the scientific viewpoint, is a transfer of information argument. Given two inaccurate but precise measurement processes, both capable of measuring the same property and each giving different values for that property, it is difficult to determine which of the values should be accepted. Statistical analysis has little to say about systematic biases and it cannot be ascertained statistically whether one or the other or both of the methods are inaccurate, the direction of the biases, the magnitude of the corrections, etc., because there is no base-line reference against which these parameters may be measured. Even greater difficulties are encountered when measurements on different properties are to be evaluated and correlated as, for example, the determination of the mass balance of

a system. Only when the measurement processes give accurate values are these difficulties overcome.

Hence, the goal of every measurement system should be one of accuracy. Further advances should not be limited by inability to measure accurately but rather by limitations in the fundamental knowledge of the system under study.

C. The Role and Limitations of a Referee Method (In General)

As was explained in section I-A, above, a referee method is a measurement method of known and proven accuracy, i.e., the systematic errors or biases of the measurement process have been found, mechanisms elucidated, and these errors eliminated, or, as is sometimes the case when actual physical elimination is impossible, corrected or accounted for. There is some misunderstanding as to the limitations and scope of a referee method:

1. A referee method is not, per se, to be a legally imposed method, the preferred method, a method to be used for qualification or certification of laboratories and the like. It may, in fact, become any of these if various organizations, assumed to have the power to so declare by law, fiat, etc., do so choose to have the referee method serve these purposes.
2. A referee method is not necessarily an easily used method; it does not have to be a method which can be used by the average practitioner in any given field. It may in fact require special skills, special equipment, controlled environment, etc., which are not readily or easily available to the average laboratory. It should not, however, be so highly specialized or unique that only a handful of scientists can use the method. This latter criterion depends to some extent on the particular field of science. In metrology there are only a small number of national laboratories

who can realize the measurement of time to an accuracy of 1 part in 10^{12}, but because this measurement at a high accuracy is readily transferred when needed elsewhere, not many principal practitioners are required. In clinical chemistry, the situation is much more complex. Here the referee method should be realizable, in the real world, by many dozens of qualified laboratories so that transfer of the proven accuracy from the referee method to the dozens of field methods may be accomplished in some reasonable time span at a reasonable expenditure of time, energy, and funds.

In sum, a referee method is a method of known accuracy, useable in the real world of measurement, one necessary link in a chain of events that leads to a meaningful measurement system.

D. Background and History

In 1967 the National Bureau of Standards (NBS) acting on the recommendations of the Standards Committees of the American Association of Clinical Chemists (AACC), the College of American Pathologists (CAP), and with the subsequent support of the National Committee of Clinical Laboratory Standards (NCCLS) initiated a program, largely with financial support from the National Institute of General Medical Sciences (NIGMS), to produce, certify, and issue Standard Reference Materials (SRM's) for use in clinical chemistry. This program has resulted to date in the issuance by NBS of 14 clinical SRM's [3]. Until 1969, NBS did not consider expanding further afield in the total clinical measurement system. However, at the International Federation of Clinical Chemistry meeting in Geneva, Switzerland in September 1969, NBS was urged by the Executive Board of that body to consider playing a coordinating role in the development of referee methods that were seen to be urgently needed. Dr. George Bowers, Jr., Hartford Hospital, and Dr.

Donald Young, National Institutes of Health, during the first half of 1970 were instrumental in helping shape the thinking and planning at NBS for the initiation of a plan whereby referee methods for clinical chemistry could be developed, tested, and validated.

Several of the subsequent participants in this work met at Hartford Hospital in May 1970 and later at NBS in June 1970 where agreement in principle was reached on the main tasks and routes to be pursued. Valuable critical advice at this stage was provided by Drs. Alan Mather and Nathan Radin, Center for Disease Control (CDC) and Dr. Roy Rand, Hospital of the University of Pennsylvania. As a result of these discussions, J. Paul Cali, NBS, prepared a memorandum that laid down, from the NBS point of view, the minimum requirements and ingredients necessary for the successful development of a referee method (Appendix 1).

The actual work began on September 10, 1970, at a meeting at NBS, when the accuracy goal for calcium in serum was set, the candidate method chosen, and various other details agreed on.

The achievement of a referee method for the determination of calcium in serum was accomplished in December 1971.

II. THEORY OF OPERATION AND PRACTICE REQUIRED FOR THE SUCCESSFUL DEVELOPMENT OF A REFEREE METHOD

A. Necessary Conditions

For the successful attainment of any Referee Method the following conditions must be met or defined:

1. An authoritative Standard Reference Material (SRM) must be available. It should be a highly homogeneous substance, usually of the highest possible purity, and its properties should be certified by a laboratory or group of the highest technical competence. (E.g., SRM's issued by National certifying bodies)
2. A method of analysis having high precision must be selected for study as the candidate for development as the Referee Method. Considerations of ease of operation, cost, and time are secondary to the achievement of a method having a known accuracy.
3. There should be a clearly defined accuracy goal for the Referee Method.
4. Laboratories of recognized technical competence must participate in the interlaboratory work.
5. The system must be in a state of quality control at all times. Quality control should be assured and maintained by properly designed and executed statistical methods under qualified and experienced statisticians.
6. A panel of qualified experts must review and direct all the technical phases of the work. As a pragmatic consideration, it is best that the coordination of the entire operation be focused in one responsible person.
7. If possible, an independent method having a known accuracy of at least a factor of 2 better than the

stated accuracy goal of the Referee Method under test should be used to measure the property(ies) under consideration, for the test samples sent to the cooperating laboratories. The question immediately occurs: If an independent, accurate method is available, why then cannot it be declared the referee method? In many instances, this does not happen because the independent method is too complex and involves specialized equipment and skills which are not readily available in the field under study. As will be made evident this situation applies in this study. (See more detailed explanation in section VIII-H.)

In the calcium Referee Method developed and reported here the above conditions were met as follows:

Under A1: NBS-SRM Calcium Carbonate (SRM 915) was available and used throughout all five exercises. The certificate of analysis for this SRM is shown as Appendix 2.

Under A2: The Experts Committee chose as the candidate method an atomic absorption method for calcium based on the work of Pybus, Feldman, and Bowers [4]. Revisions and modifications to improve accuracy were made during the course of the work and are incorporated in the final referee method that is given in full detail in section IV.

Under A3: The Experts Committee set two accuracy goals with the expectation that the more stringent, while desirable, was not likely to be attained. These were a method giving an accuracy to within ±0.5% (desirable) and ±1.0% (probable) of the true value. As work progressed, tacitly it was agreed that the latter limit was in fact the goal for this method.

Under A4: The Experts Committee chose eight cooperating laboratories to participate in the interlaboratory

work.

The names of the principal investigator and the complete addresses are given in appendix 3.

Under A5: Statistical advice, guidance, and evaluation were performed by Dr. John Mandel, Staff Statistician, Institute for Materials Research, NBS.

Under A6: The Experts Committee originally consisted of Drs. G. Bowers, N. Radin, and D. Young. Advice and criticism during the course of the work was made by Dr. J. Boutwell (CDC) and Dr. R. Eilers (CAP).

Under A7: Fortunately, at NBS an independent method for the accurate determination of calcium using an isotopic dilution mass spectrometry technique was available and used fruitfully in this work. This method served as the independent base line and is fully described in section VII, below. This part of the work was under the overall direction of William R. Shields.

B. Modus Operandi

All the necessary conditions were met and fulfilled by September 1970. The work and interrelationships of the Experts Committee (A6, above), the cooperating laboratories (A4, above), and the statistical team (A5) was coordinated at NBS through the Office of Standard Reference Materials (OSRM) under the guidance of J. Paul Cali, Chief of that office.

The actual steps followed in the work were these: (See note following)

Step 1 - Define the goals of the Referee Method in terms of what is to be achieved, use of the method, applicability, limitations, etc. (EC)

Step 2 - Set accuracy goal. (EC)

Step 3 - Provide an SRM whose purity, or other certified property, is known more accurately than the accuracy

goal set for the Referee Method. (NBS)

Step 4 - Choose an appropriate clinical analytical method as a candidate for the Referee Method. (EC)

Step 5 - Set forth the experimental and statistical designs. (EC + NBS)

Step 6 - Provide the SRM in solution and serum (or another matrix, as required) at several concentrations that are known to an accuracy greater than the accuracy goal of the Referee Method. (NBS)

Step 7 - Provide detailed work procedures for cooperating laboratories. (EC + NBS)

Step 8 - Perform laboratory experimental work using SRM solutions (for calibration purposes) and then analyze unknowns (both solutions and sera) distributed by NBS (concentrations known only by NBS). (CL)

Step 9 - Analyze results from cooperating laboratories. Elucidate and evaluate suspected biases in the candidate method. Modify procedures to eliminate suspected sources of error and to control variables within tighter limits as required. (EC + NBS)

Step 10 - Repeat steps 6 through 9 until accuracy goal is achieved. (EC + NBS + CL)

Step 11 - Publish analytical method as a Referee Method. (EC + NBS)

(Note: Groups responsible for the various steps are shown in parenthesis after each step. The following abbreviations are used: Experts Committee (EC); coordination, statistical services, SRM and sample distribution (NBS); Cooperating Clinical Laboratories (CL).

III. HISTORY OF ACTUAL WORK

A. Definitions of Terms Used

1. Exercise: The work performed by the cooperating laboratories on a specific lot of either calcium solutions or pooled sera containing calcium. Included are: (a) the instructions for that particular exercise, (b) the actual samples, (c) the data produced and analyzed, and (d) the work done at NBS on the same lot to produce values of the calcium concentration by the isotope dilution-mass spectrometric method.

Five different exercises were required. From initiation to finish, each exercise averaged two months. The various exercises are numbered from 1-5, serially.

2. Run: A complete set of measurements made by a particular laboratory on a specific, complete lot of samples in a single session of work.

In each of the five exercises, each individual sample was run in duplicate, but with the order for the entire set specified in advance. Two runs, in each of two successive weeks, were required for each set of samples for each exercise. (Exceptions will be noted in the tabular data)

3. NBS Samples: Vials or bottles containing either water solutions, calcium, or sera, and individually numbered. Calcium concentrations, determined by the isotope dilution-mass spectrometric (ID-MS) method, were not known in advance nor given to the cooperating laboratories during the course of any exercise. (One exception noted in tabular data)
4. Glassware, water, reagents, etc.: These, together with their specifications, are defined in the protocol, section IV.

B. Coding of Laboratories and Participation

Laboratories were assigned a letter. Not all laboratories participated in every exercise for reasons that are explained below. Letters assigned and exercises participated in are shown in Table 1.

Explanation for incomplete data or abort or for non-participation are as follows:

1. Lab D - Exercise 2: samples held up in shipment; several weeks storage in warm area; finally, instrument breakdown. No data used in Exercise 2.
2. Lab E - Exercise 4: Instrument instability, could not meet precision requirements; also, condition of sample led investigator to believe not able to perform quantitative transfers. No data used in Exercise 4.
3. Lab F - Exercises 4 and 5: After protocol tightened up (after Exercise 3), instrument used could not meet precision requirements. No data available for Exercises 4 or 5.
4. Lab K - Exercises 1 and 2: No participation until Exercise 3. No data available for Exercises 1 and 2.

Table 1. Laboratory Code Letters and Exercise Participation

Exercise Number	Lab Code Letters							
	A	B	C	D	E	F	H	K
1	CD	CD	CD	CD	CD	CD	CD	NP
2	CD	CD	CD	ID	CD	CD	CD	NP
3	CD	CD	CD	CD	CD	CD	CD	CD
4	CD	CD	CD	CD	ID	NP	CD	CD
5	CD	CD	CD	CD	CD	NP	CD	CD

Key: CD - Complete Data
ID - Incomplete Data or Abort
NP - Did Not Participate

C. NBS Samples--Preparation or Source, Number Coding, Calcium Concentrations (By ID-MS)

Exercise 1 - Three solutions of calcium at concentrations of (approximately) 0.08, 0.10, and 0.12 g of Ca/l were prepared at NBS using SRM 915, Calcium Carbonate. Solvents were isothermally distilled HCl and redistilled water.

Exercise 2 - Three solutions of calcium at concentrations of (approximately) 0.07, 0.10, and 0.13 g of Ca/l were prepared at NBS using SRM 915, Calcium Carbonate, in which 3.2 g of Na/l and 0.2 g of K/l were added using SRM 919, Sodium Chloride, and SRM 918, Potassium Chloride, respectively. The same solvents, as in Exercise 1, were used.

Exercise 3 - Four, one liter bottles (polyethylene) of pooled patient serum were received from Hartford Hospital. The four lots, their ID numbers (at Hartford) and approximate calcium concentrations were:

Lot 1 - serum pool mixed 26E and 27E, prepared August 1970, Ca ∿6.0 meq/l.

Lot 2 - serum pool 27A, prepared December 1970, Ca ∿4.5 meq/l.

Lot 3 - serum pool mixed 26 (B + C), prepared August 1970, Ca ∿5.0 meq/l.

Lot 4 - serum pool 26A, prepared August 1970, Ca ∿4.8 meq/l.

Exercise 4 - Same as Exercise 3, except that proportions were changed to give slightly different calcium concentrations as follows:

Lot 1 - ∿5.1 meq/l

Lot 2 - ∿5.8 meq/l

Lot 3 - ∿3.6 meq/l

Lot 4 - ∿4.1 meq/l

The sera used in Exercises 3 and 4 were prepared at Hartford Hospital according to procedures given in reference [5].

For reasons which are not entirely clear, these sera samples, after repackaging in flame-sealed vials at NBS, started to show mold and bacterial growth. It is suspected that NBS' inexperience with sterile techniques and the fairly long time between shipment from Hartford to NBS and then to the cooperating laboratories were all contributing factors. The deterioration of the serum used in Exercise 4 was, however, not evident until after repackaging and was only apparent several days after shipment to all the participants. The NBS samples for these two exercises were far from ideal and much worse than sera samples collected under usual field conditions.

In addition to the four lots of sera sent out in this exercise, an aqueous acid solution of the NBS-SRM $CaCO_3$ (SRM 915) was again prepared and sent with the four sera samples. The calcium value was ∿5.5 meq/l. It also contained sodium and potassium at the same level indicated in Exercise 2.

Exercise 5 - The sera for this exercise was prepared at CDC, Atlanta, Georgia under the supervision of Dr. David Bayse. Twenty vials of each of four concentrations and containing 50 ml of bovine pooled sera were received at NBS.

Lot 1 - ∿3.6 meq/l

Lot 2 - ∿4.3 meq/l

Lot 3 - ∿5.0 meq/l

Lot 4 - ∿5.7 meq/l

The exact procedure used in the preparation of Exercise 5 sera samples is given in Appendix 4.

All the data of this section plus the actual sample number codes and the ID-MS values are summarized in Table 2.

Table 2. NBS Samples: Composition, Source, Lot and Sample Code Numbers, and ID-MS Values

Exercise Number	Sample Composition	Source of Sample	Lot or Solution I.D.	Sample Code Numbers (inclusive)	Calcium Value (meq/l) by ID-MS
1	Ca^{++} in dilute HCl from NBS-SRM 915. No Na^{+} or K^{+}	NBS	Soln 1	1-20 and 101-120	4.0783
			Soln 2	21-40 and 61-80	5.1816
			Soln 3	41-60 and 81-100	5.8673
2	Ca^{++} in dilute HCl from NBS-SRM 915, plus Na^{+} and K^{+}	NBS	Soln 1	228-240 and 269-280	3.5928
			Soln 2	213-227 and 241-255	4.8922
			Soln 3	201-212 and 256-268	6.3767
3	Pooled patient serum	Hartford Hospital	Lot 1	301-316	6.1526
			Lot 2	321-333	4.5709
			Lot 3	341-356	5.0249
			Lot 4	361-377	4.8303
4	Same as Ex. 2	NBS	Soln 1	402-417	5.565 (Note A)
	Pooled patient serum	Hartford Hospital	Lot 1	521-535	5.0598
			Lot 2	561-576	5.7934
			Lot 3	541-555	3.6228
			Lot 4	501-516	4.1118
5	Pooled bovine serum	CDC	Lot 1	620-639	3.5704
			Lot 2	600-619	4.2939
			Lot 3	660-679	5.0234
			Lot 4	640-659	5.7325

Note A - The water solution used in Exercise 4 was not determined by ID-MS. The gravimetrically determined value was.

D. Analytical Procedures Used

The candidate referee method under development is based on the method for determining calcium in serum by atomic absorption spectrometry developed by Pybus, et al [4]. When a general method is adapted for a particular use, under specified conditions, then that particular write-up is called a procedure, or as widely used in clinical chemistry, a protocol.

Three procedures based on the Pybus, et al, method were used in the course of this work:

1. In Exercise 1, each laboratory was permitted to use the Pybus method (as if it were a detailed procedure) as published in Clinical Chemistry. Specific details not expressly stated were left to the discretion of each participant. Significant departures were to be reported. Calibration of the atomic absorption spectrometer was to be by an accepted procedure using a calibration curve and the procedure used was also to be reported. Finally, any deviation from reagent specifications was to be reported.
2. In Exercises 2 and 3, the Pybus method was amended to provide more specific instructions in several areas. In addition, clarification of several points was made. The first revised procedure will not be given here. The revisions included: (a) a glossary of terms used, which had caused some confusion in Exercise 1, (b) a brief explanation of how the ID-MS method would be used to provide the reference base-line against which the interlaboratory results would be measured, (c) more detailed instructions on the preparation of the various solutions used, (d) a clarification of the section specifying the order in which the blank, standard solutions and unknowns are to be determined, and (e) a modified reporting procedure so that all ten readings of the absorbance

for a given unknown would be recorded.

3. In Exercises 4 and 5, a further revision was made, following a meeting at NBS of all participants, to tighten up several steps. These included: (a) more rigid pipetting instructions, including a technique to insure good rinsing while saving sample solution, (b) an exposition of what constituted a valid measurement and explicit rules for discarding data, (c) an exposition of the procedure to be followed in the determination, (d) reporting of significant figure, (e) importance of temperature control, (f) re-emphasis on quality of water used and specifications for reagents, and (g) emphasis on assuring the stability of the spectrometer.
4. The final procedure, rewritten for this publication, is given in section IV.

E. Reporting of Data

Each cooperating laboratory received with the actual samples, data sheets for recording all pertinent data. The forms used are shown in Appendix 5. In addition, comments regarding deviations from the protocol, troubles encountered, suggestions for improvement, etc., were returned to NBS with the completed data sheets for each exercise.

IV. REFEREE METHOD - SCOPE, LIMITATIONS AND PROCEDURE

A. Scope

The procedure given in Section C, below, is derived and adapted from the candidate method of Pybus, et al [4]. The referee method when used exactly as given, subject to the limitations stated, will give the concentration of calcium in serum with an accuracy within ±2% of the true value.

The referee method as given here is not to be considered suitable for routine use in the clinical laboratory. The size of sample required, precautions that must be adhered to, time required, and cost preclude its use under normal operating conditions. This should not be construed to mean that this method cannot be adapted for routine use as indeed was the case in the original Pybus, et al., method. Its primary utility lies in the following:

1. To establish the absolute value of calcium in control or pooled sera.
2. To establish the validity of calcium values given in various clinical kits.
3. To be a method against which the many field methods for calcium may be tested to ascertain the accuracy of those methods.
4. To be used by manufacturers of calcium reference materials, kits, control sera, etc., in establishing and maintaining the quality of these products.

B. Limitations

During the course of these studies, several important limitations became evident. The following conditions must be met to attain the accuracy inherent in the referee method:

1. The procedure must be adhered to in all details. Deviations, short-cuts, adaptations, etc., are not allowed.
2. All glassware involved in volumetric measurements,

i.e., pipets and volumetric flasks, must meet NBS Class A specifications. These items are commercially available.

3. Reagents, including water, must meet specifications set forth in the procedure.
4. In this work three different spectrometers were used (see Table 3). All are double-beam instruments. It is not possible to say whether or not other instruments can provide the necessary stability, precision, etc., to accomplish the obtainable accuracy of the referee method. In one instance, Lab F dropped out of Exercises 4 and 5 when it was determined that its instrument could not meet the precision requirements set forth. It is clear that the instrument used must be in optimum operating condition, highly stable, and linear throughout the range of interest. Flame conditions used are shown in Table 4.
5. The work must be performed by workers well acquainted with the techniques of analytical chemistry. Weighings, aliquoting, preparation of standard, blank, and unknown solutions must be performed with the highest regard to accuracy in order to limit errors from these steps to much less than 1%, if the overall accuracy of the method is to be achieved.
6. Sufficient, uninterrupted time ($\sim$1 week) must be made available to perform the method. It cannot be worked into the daily work load of the clinical laboratory. Finally, the worker who performs the method must be highly motivated. This implies a scientist who enjoys working meticulously and wants to provide analytical data that will stand up to the most critical analysis.

Table 3. Instruments Used and Operating Conditions

Lab Code	Instrument Company/Model	Fuel	Flame Condition	Source Current	Readout	Ambient Temperature Range
A	P&E - 403	Air - C_2H_2	stoichiometric	15mA	Digital	23°C
B	IL - 153	Air - C_2H_2	reducing	7mA	Digital	20-26°C
C	P&E - 303	Air - C_2H_2	sl. reducing	10mA	Digital	23-25°C
D	P&E - 303	Air - C_2H_2	oxidizing	15mA	Recorder	21-27°C
E	P&E - 303	Air - C_2H_2	reducing	10mA	Digital	22-27°C
F	P&E - 303	not given	reducing	15mA	Rotating Dial	25.8-29.5°C
H	IL - 153	not given	stoichiometric	8mA	Digital	23-24°C
K	IL - 153	Air - C_2H_2	stoichiometric	8mA	Digital	21-26°C

Note: The Burner used in all laboratories was a Boiling Triple Slot Burner.

Table 4. Flame Conditions, Pressure and Flow Rates

Lab Code	Air		Acetylene	
	PSI	l/min	PSI	l/min
A	30	21.2	10	8-9
B	37-38	8.8	5.4-6	1.7
C	30	16[a/]	8-8.5	4
D	30	7.5	8-8.5	9
E	8.5	NR	9	NR
F	28	NR	8.5	NR
H	14	NR	4	NR
K	5	NR	5.5	NR

NR - Not Reported
a/ - Exercise 2 only 7.5 l/min

C. Procedure

Samples

The method is designed to provide accurate values of the calcium concentration of serum or aqueous solutions as provided for analysis. The method does not assure adequacy or integrity of sampling per se, but if a representative sample is taken, the method will provide an accurate value of the calcium in that specimen. Whether the specimen is truly representative, or homogeneous, is beyond the scope of this method. The only requirement is that a representative sample of the presented specimen be pipetted accurately. The method has applicability, at the specified accuracy, in the range from 3 to 7 meq Ca/l.

Reagent Specifications

1. Water, preferably distilled and deionized, should measure at least 10^6 ohms specific resistance at 25 °C. It should be available in large quantity for use as a diluent and for the final rinse operation on all glassware and apparatus coming in contact with the solutions involved. Only water that meets these specifications is to be used in these operations.
2. Calcium standard solutions should be prepared from $CaCO_3$ issued and certified by NBS. Its identification number is SRM 915. This material should be dried for 4 hours at 200 °C and cooled to room temperature in a desiccator before use.
3. Lanthanum oxide should be of high purity and known to contain less than ∿15 μg of calcium per gram.
4. Sodium, potassium, and strontium chlorides should be ACS Analytical Reagent Grade (AR) quality.
5. Hydrochloric acid, meeting ACS-AR specifications should be used.

Glassware Specifications

All glassware required--10 ml volumetric pipet (to contain); 500 ml volumetric flasks--should meet NBS Class A specifications.

All glass or plastic surfaces coming into contact with reagents, water, diluent, or sample must have been previously cleaned as follows:

1. Use routine cleaning procedure (hot water with detergents, plus usual rinses)
2. Soak glassware for 30 minutes in 0.05 M HCl.
3. Rinse with several portions of distilled water, (5-6 minimum)
4. Air dry (inverted) in a dust-free environment.

Preparation of Reagents

1. Stock Blank Solution (140 mmol in NaCl and 5.0 mmol in KCl per liter). To a clean 1-l volumetric flask, add 8.18 g of NaCl and 373 mg of KCl. Dissolve in H_2O and fill to the neck. When at working temperature (ambient), dilute to calibrated volume and mix by inverting the flask 30 times.
2. Diluent Solution (10 mmol $LaCl_3$ and 50 mmol HCl per liter). (Plan to make sufficient diluent for the work to be performed in one continuous series.) Transfer 1.63 g of La_2O_3 to a 1-l flask and dissolve in 10 ml H_2O and 6.7 ml concentrated HCl. After the La_2O_3 is dissolved, dilute with water to the neck of the flask. When the solution has reached ambient temperature, dilute to calibrated volume and mix by inverting the flask 30 times. (NOTE: If an internal reference is to be used, add 30.6 mg $SrCl_2 \cdot 6H_2O$ per liter)
3. Standard Stock Solutions of Calcium. Prepare a minimum of three concentrations at 4.00, 5.00, and 6.00 meq Ca/l with each to contain 140 mmol NaCl

and 5.0 mmol KCl per liter. To each of three 1-liter volumetric flasks, add 8.18 g NaCl and 373 mg KCl. To the first flask (4.00 meq Ca/l) add 200.2 mg of $CaCO_3$, to the second flask (5.00 meq Ca/l) add 250.2 mg $CaCO_3$, and to the third flask (6.00 meq Ca/l) add 300.2 mg $CaCO_3$. (Use only NBS-SRM 915 previously dried at 200 °C.) To each flask add a few ml of H_2O and 1 ml concentrated HCl. Make sure all the $CaCO_3$ is in solution before diluting with H_2O to the neck. When at ambient temperature, dilute each flask to calibrated volume and mix by inverting the flask 30 times. Label all flasks appropriately.

Procedure for Diluting

All solutions should be at ambient temperature.

All solutions, except the diluent, but including the unknown samples, will be diluted to 1:50 using a 10 ml volumetric (TC) pipet and 500 ml volumetric flasks. Only one 10 ml volumetric pipet is to be used throughout to reduce errors caused by differences in drainage times between the aqueous or dilute acid and sera solutions.

1. Transfer to a 500 ml volumetric flask approximately 450 ml of diluent stock solution. Add to the flask 10.00 ml of the blank stock solution using the 10 ml pipet. (NOTE: The tip of the pipet should remain in the diluent to prevent the formation of foam by the serum) After drainage of the pipet stops, using a rubber bulb, gently blow out the residual liquid. Rinse the pipet three times with diluent from the flask, each time returning the pipet contents to the flask by drainage and blowing.
2. Dilute to calibrated volume with diluent and mix thoroughly with 30 inversions. Set aside.
3. Draw into the pipet H_2O contained in a clean

beaker. Fill to slightly above the mark and discard.

4. To condition the pipet, fill to 1-2 mm above the mark with 4.00 meq Ca/l standard stock solution. Discard. Repeat twice.
5. To a 500 ml volumetric flask, transfer 10.00 ml of the 4.00 meq Ca/l standard stock solution using the technique described in steps 1, 2, and 3, above.
6. Repeat step 5, twice, but using the 5.00 and 6.00 meq Ca/l standard stock solutions. Condition the pipet each time as given in step 4, using the approximate standard stock solutions.
7. After the blank and standard solutions have been diluted and the pipet rinsed with H_2O, draw 2-3 ml of the first unknown solution into the pipet. Place a finger over the end of the pipet and withdraw from unknown solution container. Tilt the pipet to a horizontal position and slowly rotate the pipet to wet thoroughly all internal surfaces. Allow a small amount of air to leak past the finger so that the rinse solution may come into contact (a small way above the mark) with the upper stem surface. Discard. Repeat the rinse and conditioning operation once more.
8. Fill the pipet to the mark with the unknown solution and deliver into a clean 500 ml volumetric flask. Rinse the pipet three times with diluent solution, returning all rinses to the flask. Dilute the flask to calibrated volume with diluent solution and mix by inverting 30 times.
9. Rinse the pipet with water, condition with the next unknown solution, and repeat steps 7 and 8 as many times as there are unknowns to be analyzed. (NOTE: If the rinse solution doesn't

completely wet the sides of the pipet, clean the pipet by rinsing it several times in a solution of aqua regia ($HCl-HNO_3$ 3:1). Then, repeat step 9.)

10. At the conclusion of the dilution procedures there should be:

    a. One 500 ml volumetric flask containing a 1:50 dilution of the blank stock solution. Label "B".
    b. Three 500 ml volumetric flasks each containing 1:50 dilutions of the calcium stock standard solutions. Label "4.00", "5.00", and "6.00" (these are in meq Ca/l).
    c. As many 500 ml volumetric flasks each containing a 1:50 dilution of each of the unknown solutions as there are to be analyzed. Label appropriately. (NOTE: For the sake of clarity in following steps it is assumed there is one unknown labeled "x".)

Atomic Absorption Spectrometry (AAS) Measurement Procedure

It is assumed that the operator is fully familiar with the instrument to be used. It is not possible in this method to give detailed instructions necessary to assure instrument stability, linearity, flame conditions, etc. In general, the accuracy of the method cannot be attained unless the instrument is in optimum operating condition and meets all the specifications set forth by the manufacturer. Repeatability of readings of the same solution within ±0.5% (maximum) is a necessary condition.

1. Instrument and Electrical Adjustment. Prepare the atomic absorption spectrometer for operation according to instructions provided in the operator's manual. Place the calcium hollow cathode lamp in the lamp housing receptacle. Turn the power supply switch to "On". Select the optimum current for the lamp, and allow ample "warm-up" time for the lamp to become stable. Adjust the monochromator slit and set the wavelength selector to the calcium resonance line at 422.7 nm or 4227 Å. Adjust the photomultiplier dynode voltage to give optimum current output with minimum dark current.
2. Flame Condition. Open the tank valves on the air and acetylene supplies. Adjust the secondary regulators as recommended by manufacturer. Check the burner to make sure the premixing chamber and nebulizer are clean and free of any foreign obstructions. Insert a three-slot (Boling) burner head on the burner. Light the burner and adjust the air and acetylene flow rates recommended for the instrument. To stabilize the temperature of the burner head, aspirate water into the flame for at least 10 minutes before proceeding to the next step. (NOTE: A fuel-rich air-acetylene flame gives optimum sensitivity for the measurement of calcium; however,

it may be difficult to obtain the precision specified in this method with a fuel-rich flame. Therefore, it is suggested that a stoichiometric or slightly fuel - rich flame be used to obtain the highest precision for calcium in serum.)

3. Determination of Optimum Absorption. Determine the stability and repeatability of the instrument as well as the calibration curve as follows:
   a. Adjust the instrument to zero absorbance while nebulizing water.
   b. Nebulize the solution of 4.00 meq Ca/l and measure the absorbance.
   c. If the absorbance is not 1.000 units, adjust the scale expansion of the readout system until the absorbance value is greater than 1.000. (NOTE: A scale expansion of approximately 5 is required for most instruments.)
   d. Readjust the instrument to zero absorbance with water.
   e. Nebulize the reagent blank, 4.00, 5.00, and 6.00 meq Ca/l and record their absorbances. Nebulize water between each of the standard solutions and check the zero value.
   f. Repeat the sequence of blank and standards as outlined in 3e until a repeatability of readings for the same solution is within ±0.5%
   g. Subtract the absorbance value for the reagent blank from the average value obtained for the standard solution.
   h. Plot on rectilinear graph paper the absorbance, corrected for reagent blank, as ordinate, versus the concentration of standard expressed in meq/l. A typical calibration curve is shown in Figure 1 of Ca as abscissa. (NOTE: If concentration values are determined directly from the instrument

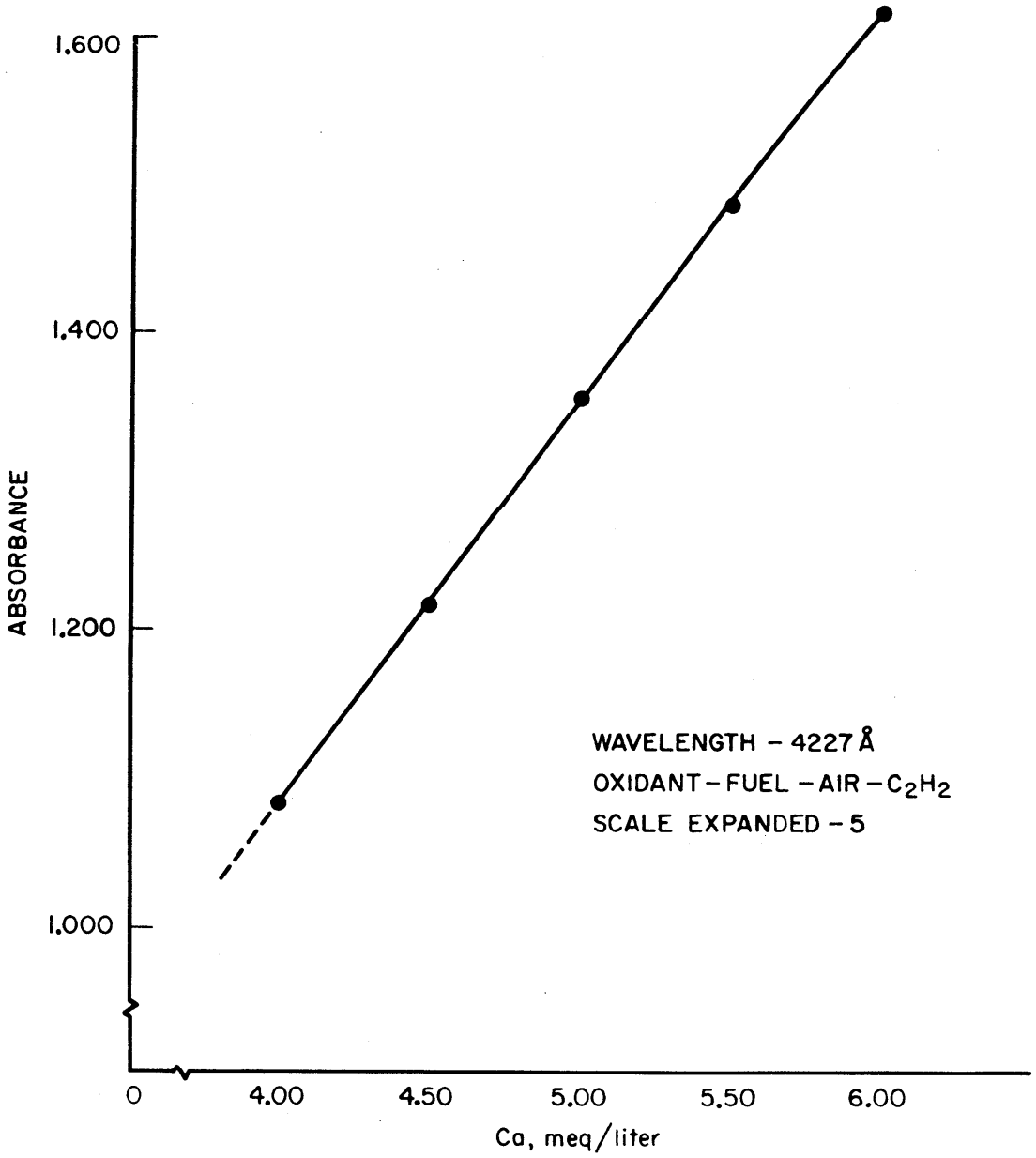

Figure 1. Typical Calibration Curve for the Atomic Absorption Spectrometry Determination of Calcium

instead of the absorbance values, follow the same procedure.)

i. If the calibration curve is not linear, prepare calcium standards of 4.50 and 5.50 meq/l following the procedure given previously.

4. Absorbance Measurements. Measure the absorbance of the unknown solution as follows:
   a. Repeat the calibration curve as outlined in Step 3.
   b. Nebulize the unknown solution and then nebulize the two standard solutions that are closest to the value of the unknown. Record these absorbance values.
   c. Repeat this sequence of standards and a single unknown until 10 valid measurements have been obtained.
   d. Repeat steps 4a-4c on additional unknowns and their associated standard solutions.

5. Valid Measurement. To obtain a valid measurement, follow the sequence of standards and unknowns and record the data as follows:

Test No.	Ca meq/l Sample	4.00	4.50	5.00	5.50	6.00
1	Standards	1.084	1.222	1.355	1.486	1.608
2	Unknown (1)			1.477		
3	Standards			1.354	1.480	
4	Unknown (2)			1.475		
5	Standards			1.351	1.488	
6	Unknown (3)			1.420		
				1.320	1.475	

Two valid measurements for the unknown in test 2 and 4 were obtained as the difference between consecutive standards is less than 1%. However, the value obtained in test 6 is not valid because the difference between 1.351 and 1.320 is greater than 1% E.g.

1.355 - 1.354 = .001 or <.1%
1.486 - 1.480 = .006 or .4%
1.354 - 1.351 = .003 or <.1%
1.488 - 1.480 = .008 or .5%

6. Calculations. Calculate the sample concentration by an interpolation technique.

   a. Mathematical Interpolation. Calculate the concentration in meq Ca/l for each valid measurement using the following formula:

$$C = S_1 + \frac{(A_x - A_{s_1})}{(A_{s_2} - A_{s_1})}(S_2 - S_1)$$

where C = sample concentration; meq Ca/l
S_1 = concentration of lower standard, meq Ca/l
S_2 = concentration of upper standard, meq Ca/l
A_x = absorbance of unknown
A_{s_1} = absorbance of lower standard
A_{s_2} = absorbance of upper standard

   b. Graphic Interpolation. Plot the absorbance values for the upper and lower standards versus concentration meq Ca/l on rectilinear graph paper for each set of data. Draw a straight line between data points and determine the concentration of the unknown from the curve.

   c. Computer Interpolation. By means of a least square plot using the paper and lower standards, compute the concentration of the unknown with a computer.

   d. Compile the concentrations obtained by method a, b, or c for the ten valid measurements and determine the mean value. For the best precision methods 6 a or 6c are recommended.

V. COMMENTS OF THE COOPERATING LABORATORIES

During the course of the five exercises and especially at the meeting held at NBS in June 1971, the cooperating laboratories made comments, suggestions, and criticisms. It is not necessary to list each of these, but those which had a bearing on the future exercises, or were adopted, or threw light on current steps in the procedures are given below:

Lab A - Exercise 4 - Serum samples for this exercise contained sufficient insoluble residue so that reproducibility of aliquoting was seriously impaired. Precision on separate aliquots taken 6 days apart ranged from 0.3 to 1.7%.

Lab A - Exercise 5 - Only deviation required from the recommended procedure was to rinse the pipet with a solution of $HCl-HNO_3$ (3:1) to assure proper drainage.

Lab B - Exercise 1 - Adequacy of rinsing the pipet according to the procedure was questioned. The greatest source of instrument error was thought to be in making the reading. Because the published Pybus, et al., method was designed to measure Ca in the presence of Na and K, these were added to the NBS unknown samples. Auxiliary work showed a 1% lower result when Na and K were absent. Suggested that in a future exercise test samples be supplied with and without Na and K to verify this finding.

Lab B - Exercise 2 - $Mg(C_2H_3O_2)_2$ added to standard solution, and internal reference to SRM used throughout.

Lab B - Exercise 4 - Two days work required to prepare glassware and make up solutions.

Lab B - Exercise 5 (First set of data not used and rerun made.) - Method run by senior technologist and not professional chemists. Second run 1-2% higher than run 1. Significant unauthorized changes made in

procedure; e.g., three 5 ml pipets used instead of one 10 ml pipet; dilutions made were 5:250 instead of 10:500; ambient temperature conditions not controlled. Furthermore, the technologist was not enthusiastic about performing the work.

Work rerun by willing senior personnel who followed protocol exactly.

Lab C - Exercise 1 - Protocol followed as faithfully as possible. Instrument problems experienced over past 3 weeks; occasionally high-noise level was observed. A rather large variation between sets of samples in run 2 noted. It was difficult to adjust instrument in mid-run and results were calculated relative to the immediately preceding 5 meq/l standard.

Dilution technique was changed. A single class A 50 ml volumetric pipet was used to deliver all the diluent and a single class A 1 ml pipet used for standards and samples. The 1:50 dilution does not give satisfactorily high absorbance values to feel we are performing with the highest accuracy and precision of our instrumentation.

La_2O_3 and $Mg(C_2H_3O_2)_2$ used were from different suppliers than those cited in original candidate method.

Glassware was washed according to routine procedures prevailing in this lab: (1) soak for 24 h in chromic acid, (2) rinse 4 times with deionized H_2O, (3) rinse 4 times with deionized distilled H_2O, and (4) air dry under cover.

Lab C - Exercise 3 - Absorbance and concentration values were based on the average of 10 readings.

Because of insufficient sample received, 3 samples were diluted 5:250 instead of 10:500.

The sera samples for this exercise with apparently actively-growing cultures led to some difficulties

with pipetting. If large dilution volumes are to be maintained, an anti-foaming agent should be used in the future.

As a result of discussions at the meeting held at NBS, June 1971, there is some confusion concerning the directions of the present protocol. Suggest it be extensively modified and clarified to pin down exact procedures to be used concerning "valid measurements" and rules for discarding data. Also the order in which the sample, standard, blank, are to be run is not exactly clear in the present protocol.

Lab C - Exercise 4 - This lot of sera even worse than those for exercise 3; sampling errors due to the growth are possible.

In run 1--2,0,0,1 sets of measurements rejected as not meeting the valid measurement criteria. Then 5 sets rejected for the next sample, so the aspirator was cleaned and a new sequence started with zero subsequent rejections. The rest of run 1 saw 1,2,2,1 and 2 sets rejected for the sample sequences remaining.

In run 2, for the five NBS samples run in duplicate, the following sequence of rejected readings was obtained: 1,3,1,7,3,1 entire set of 10 readings, 2,7,2, and 6 respectively, for the corresponding samples sequences.

Lab D - Exercise 1 - Dilution procedure used 1 ml pipet to 50 ml volumetric flasks. Calibration curves are not exactly linear over the range 4-6 meq/l. Results are calculated from a calibration curve constructed for each batch. Curve shows small variations from batch to batch. Standards for calibration are run in duplicate followed by quality control specimens. Patient specimens are run in groups of five followed by 2 standards used as drift markers. Changes in these outside $\pm$ 1% is the criterion for recalibration.

No Sr internal standard used. No blank solutions used since measurements made a relative basis, i.e., sample absorbance is made relative to absorbance of calibrating standards.

Lab D - Exercise 2 - Protocol calls for a double rinsing of pipet and 10 ml sample size. With 2 runs to be made there is insufficient sample. Suggest a partial filling of pipet (for rinses) and then inversion.

Hard to follow protocol, especially as to deadlines (schedules) because the work must be accommodated to an extremely busy analytical system.

Samples received for exercise 3 not considered fit for analysis; had sediment and floating particles.

Ten replicate readings not provided as normal procedure is to run at steady state, which is 2-3 minutes. The steady state achieved (shown on recorder chart attached) indicates the variation that would be expected if interrupted replicate determinations had been made.

Lab D - Exercise 4 - Due to routine work load, repeated sets as described in the protocol were not able to be run. Duplicate readings provided from chart recorder readout system.

In both rounds 1 and 2, two sets of data were rejected.

Lab D - Exercise 5 - New burner assembly unit installed. Calibration curve is now linear.

Lab E - Exercise 1 - Reagents required by the protocal not arrived. Ran exercise using our present method based on the Trudeau and Freier method [6].

Lab E - Exercise 4 - (Exercise aborted)-Quantitative transfer of turbid and contaminated solution was most difficult due to clogging of the pipet and the flocculent debris that adhered to the inner surfaces.

Visible particles observed in the final diluted solutions. Protocol followed exactly but instrumental precision required by the protocol could not be obtained. Decided to abort remainder of exercise.

Physical and instrumental problems should be alleviated when we move to new quarters.

Lab F - Exercise 1 - Instrument location is not ideal; flame is subject to drafts. Attainment of stable meter readings has always been something of a problem. For each run, standards from lowest to highest concentration were run, followed by the unknowns in assigned order. Baseline was set with diluent containing La and Sr and checked and reset, if necessary, after at least two sample readings.

Readout indicator was calibrated in percent absorption and read to nearest hundredth although movement of the dial several units in either direction often had no noticeable effect on the meter needle.

Graphical plot of points on calibration curve was not precisely linear, and best fit standard curve was computed. Values of unknowns for three ranges of standards were computed directly, and were compared with values taken directly from the calibration curve. Differences of 1-3% were noted. This raises the question of what is the proper procedure for calculating the unknowns.

Unknowns and standards were prepared the day before the run was made. A 1 ml pipet was used and rinsed with deionized H_2O rather than saline. No particular preparation of the instrument made, e.g., burner head not cleaned. A different batch of diluent prepared for run 2.

Lab F - Exercise 2 - Protocol for standard values bracketing the unknown to be within 1% of each other. The

maximum difference between standard readings over a series of 10 determinations (in a set on one unknown) was 3.1%.

Each pair of standard readings used to bracket the unknown was averaged and the average of the 10 readings used to compute the reported value.

Lab F - Exercise 3 - In order to verify matrix effects, both $Mg(C_2H_3O_2)_2$ and $SrCl_2$ were omitted from diluent and stock standard solutions. Slope of the standard curve fell from ∿45° to one-half that.

Deviation from the protocol is noted. It was difficult to maintain the % absorption values for the standard over the 5-hour period of the run. With care this could have been done, but would have doubled the time devoted to the exercise. Nevertheless, the two standard values used to bracket the unknowns were always reproduced to approximately ±1% of each other. The method of calculation was thus modified from that of picking unknown values from a standard curve to that of ratioing the unknown reading to that of the average reading of the bracketing standards and multiplying this factor by the known standard value.

Are other labs having difficulty in maintaining a stable standard curve throughout the run?

Lab F - Exercise 4 Our instrument has no better than a 30% chance of any pair of readings reproducing within ±0.5% (revised protocol requirements). We will pass for this exercise.

Lab H - Exercise 5 - The need to rerun Exercise 5 has led us to investigate our glassware washing procedures. We found some white residues on several volumetric flasks. Tap water may still be in some flasks before drying, since deionized water rinsing is a hand operation. For the rerun we pre-rinsed flasks with deionized water and diluent before use.

Within run precision is better than run-to-run agreement.

Lab K - Exercise 3 - Followed by Pybus, et al., method, except no integration used; internal standard used. Instrument extremely stable; readings obtained directly. Unable to complete every aspect of this exercise (due to time limitations) and values reported are not based on 10 independent readings.

Lab K - Exercise 4 - Instrument overhauled after completion of Exercise 4. Found some dirt-film on internal optical surfaces. After cleaning sensitivity increased and noise level decreased.

Sample #564 had fibrin clots.

VI. RESULTS

(John MANDEL)

A. Introduction

The data in this report cover five separate "exercises", each of which is an interlaboratory study conducted at 3 to 5 levels of concentration. In the course of running these five exercises, a number of modifications were introduced, both in the basic protocol and by the individual laboratories in their equipment, personnel and procedure. Furthermore, not all laboratories participated in all five exercises. These circumstances make it undesirable to run a single overall analysis covering all five exercises. Instead, individual analyses were run on the data for each exercise, and subsequently, the data for each laboratory were studied separately over the period covering all five exercises. The objectives of the statistical analysis were as follows:

1. Find measures for within and between laboratory variability for each exercise.
2. Determine deviations from the "target values" (in this case, the values given by the isotope dilution method) for all measurements by all laboratories, Study the pattern of these deviations, both between laboratories within each exercise, and between exercises, for each laboratory.

The first objective covers precision; the second, essentially accuracy.

B. Summary of Precision and Accuracy Data

Table 5 is a summary showing, for each solution in each exercise, the variability between duplicates (same laboratory, same run), between runs (same laboratory), and between laboratories. The table also shows the bias, i.e., the difference between the average value for all laboratories and the isotope dilution value, for each solution, expressed as percent of the

isotope dilution value.

The precision data are to be interpreted as follows: For "replicates", the statistical population considered is a hypothetically infinite set of replicate values obtained by any given laboratory in a single run. For "runs", the population is a hypothetically infinite set of values made up of single determinations, all from the same laboratory, but each from a different run. For "labs", the population is a hypothetically infinite set of values made up of single determinations, each from a different laboratory. In each case, the measure (%CV) is the estimated standard deviation for the population considered, expressed as percent of the isotope dilution value.

A more detailed presentation of the data is afforded by the graphs shown in Figures 2 to 16. These will be discussed further in this report, following the discussion of the Summary Table.

C. Discussion of Summary Table

From the last column of Table 5 it is apparent that on the whole, the atomic absorption method, as carried out by the laboratories participating in this study, gave slightly lower values than the isotope dilution method. Only two of the nineteen bias values are positive, and both of these are small. It is also apparent that with the exception of Exercises 3 and 4, both of which used questionable serum solutions, the bias is generally small compared with the variability between laboratories. Thus, while a small negative bias appears to be present in the atomic absorption method, this bias is essentially negligible compared with the larger random errors between laboratories.

The summary also shows a noticeable effect of runs, especially for the serum solutions (Exercises 3, 4, and 5). Such a situation is common in analytical chemistry: measurements made at several days' interval generally show less agreement than measurements made in replicate at essentially the same time.

Finally, laboratories show measurable systematic differences from each other, as evidenced by the fact that the coefficient of variation between laboratories (Table 5) is practically always larger than that between runs. Again this effect is slight for aqueous solutions (Exercises 1 and 2), but more noticeable for serum solutions (Exercises 3, 4, and 5).

In evaluating the summary data, it must be kept in mind that the coefficient of variation is related to a one-standard-deviation effect. One can expect a range of ±1.96 standard deviations to cover approximately 95% of all measurements. Thus, to cover 95% of the population of measurements in each category, the coefficients of variation in Table 5 should be multiplied by 1.96 and allowance should be made for that variation in both directions.

Elsewhere in this report, the vicissitudes of the participating laboratories in the course of this study are reported in detail. From these it follows that reasonably satisfactory and stable conditions, both in the sample preparation (of serum) and in the laboratories' techniques, were not achieved until Exercise 5. The data in the summary table show that even then, as much as 2 to 4% overall range of variation over the various laboratories is likely to occur.

D. Graphical Report of Results

Figures 2 to 6 show the individual results for all participating laboratories for each of the five exercises. The plotted values are the deviations of individual measurements* from the corresponding isotope dilution values. For each laboratory, ten values for run 1 are plotted first for all solutions, followed by the values for run 2.

Figures 7 to 14 contain the same data, but each figure now represents a single laboratory and the values are plotted in order of successive exercises.

For purposes of comparison, the results of the isotope dilution method are presented in similar graphical form in Figures 15 and 16 in which each point represents a single determination. For Figure 15, dealing with the aqueous solutions, the deviations are calculated with respect to the value provided by the known weights of calcium, salt and water. In Figure 16, dealing with serum solutions, the deviations are calculated with respect to the average value of the isotope dilution determinations. Thus, Figure 16 provides information solely on the precision, whereas Figure 15 provides information also on the accuracy of the isotope dilution method.

Figures 7 to 14 essentially confirm the conclusions stated

---

*except for Exercise 1, in which averages of several measurements are plotted. The design for Exercise 1 was slightly different.

above. In addition, they allow us to follow the events as they occurred in the various laboratories in the course of this study. If we exclude Exercises 3 and 4, because of the unsatisfactory condition of the serum, we find that only 4 laboratories (A, B, C, and E) stayed consistently within a band of ±2% deviation from the isotope dilution value. With great care, both in the preparation of the sample, and in the laboratory technique, it appears that a goal of ±2% overall error could be achieved. The data also show that any significant departure from these conditions of constant concern and control will result in much larger errors.

TABLE 5. SUMMARY OF PRECISION AND ACCURACY RESULTS

Exercise	Solution No.	Isotope Dilution Value	%C.V. Replic.[a]	%C.V. Runs[b]	%C.V. Labs[c]	% Bias[d]
1	1	4.0783	1.1	1.1	1.1	- .02
	2	5.1816	.7	.7	1.0	- .30
	3	5.8673	.9	1.0	1.2	- .76
2	1	3.5928	1.3	1.4	1.7	.09
	2	4.8922	.8	1.3	1.7	- .10
	3	6.3767	.9	1.4	1.5	- .63
3	2	4.5709	.4	.8	2.7	- .79
	4	4.8303	.8	1.2	4.0	-1.62
	3	5.0249	.4	.6	3.9	-1.75
	1	6.1526	.3	1.0	2.2	-2.28
4	3	3.6228	.6	1.4	1.6	- .15
	1	4.1118	.2	.6	1.1	- .17
	2	5.0598	.2	.4	1.0	- .93
	5	5.5650	.2	.4	.9	-1.20
	4	5.7934	.3	.5	1.6	- .93
5	1	3.5704	.3	1.3	2.2	.24
	2	4.2939	.5	1.0	1.4	- .48
	3	5.0234	.2	.5	.7	- .13
	4	5.7325	.2	.4	.8	- .18

(a) Percent coefficient of variation for replicates within runs.

(b) Percent coefficient of variation for determinations from different runs.

(c) Percent coefficient of variation for determinations from different laboratories.

(d) Difference between average of all labs and ID-MS value, as % of ID-MS value.

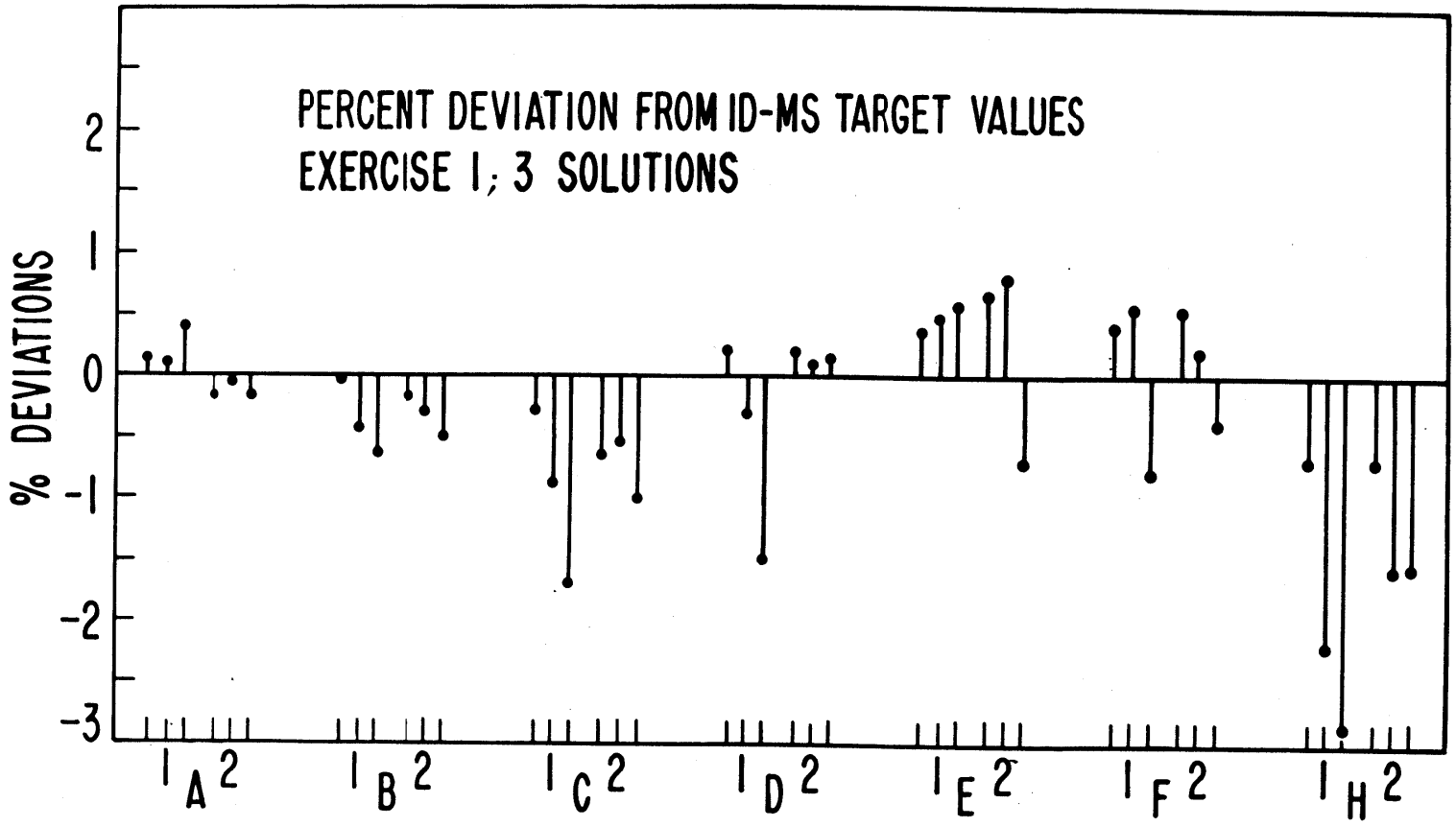

Figure 2. Laboratories Code and Round Numbers

[NOTE: Figures 4 and 6 have been photographically reduced to fit space requirements. Therefore, visual comparisons among the figures should be avoided.]

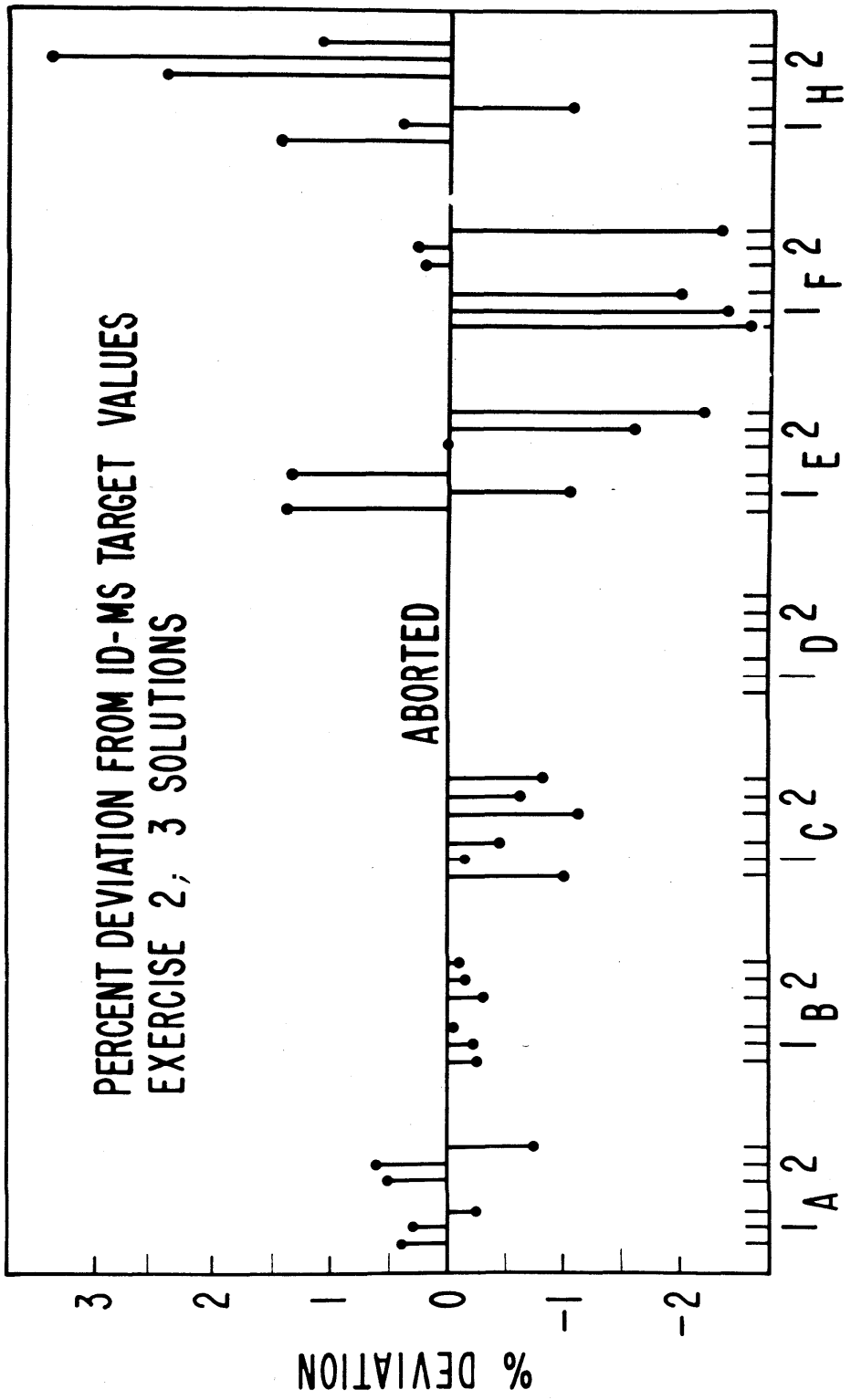

Figure 3. Laboratories Code and Round Numbers

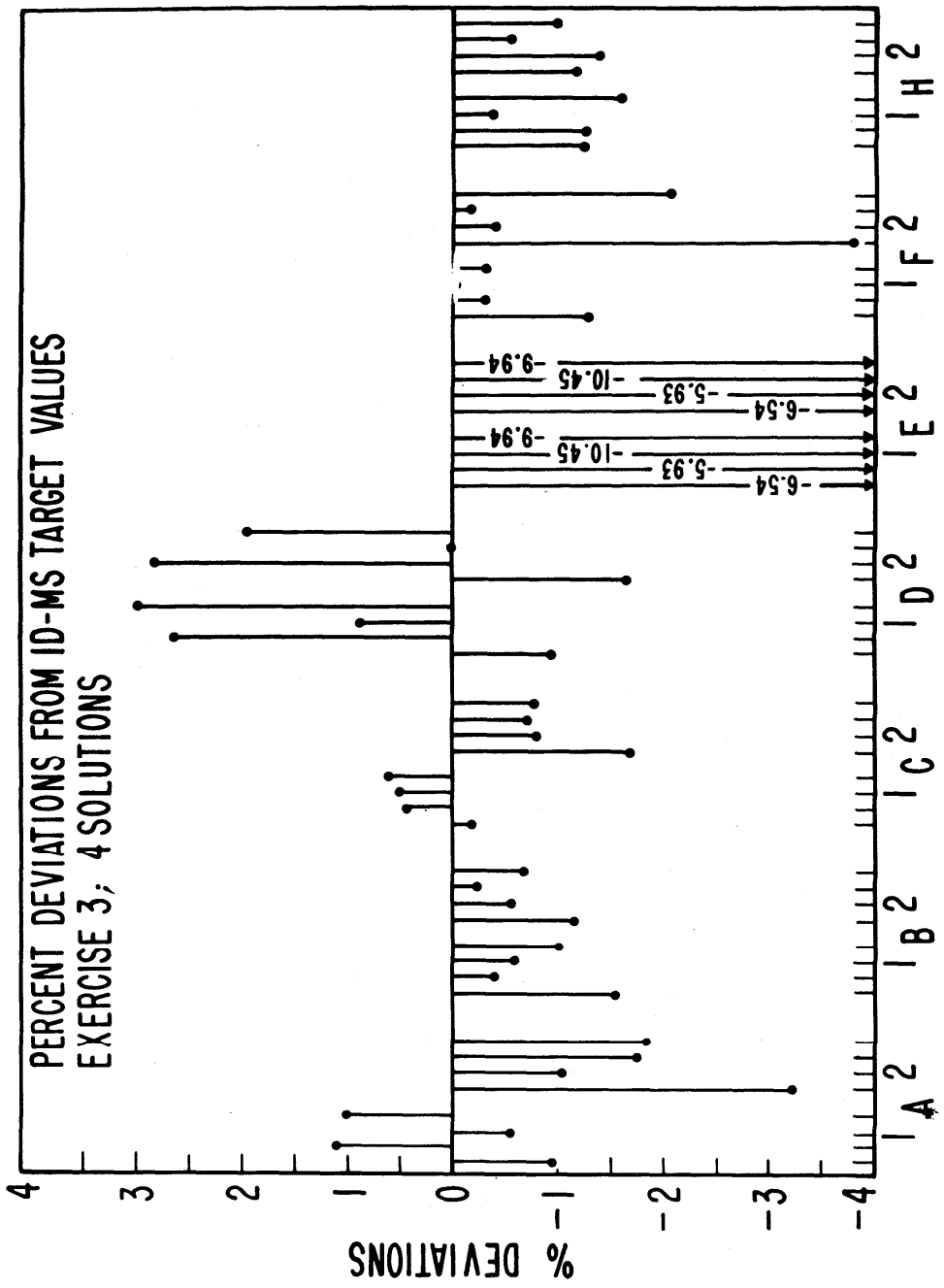

Figure 4. Laboratories Code and Round Numbers

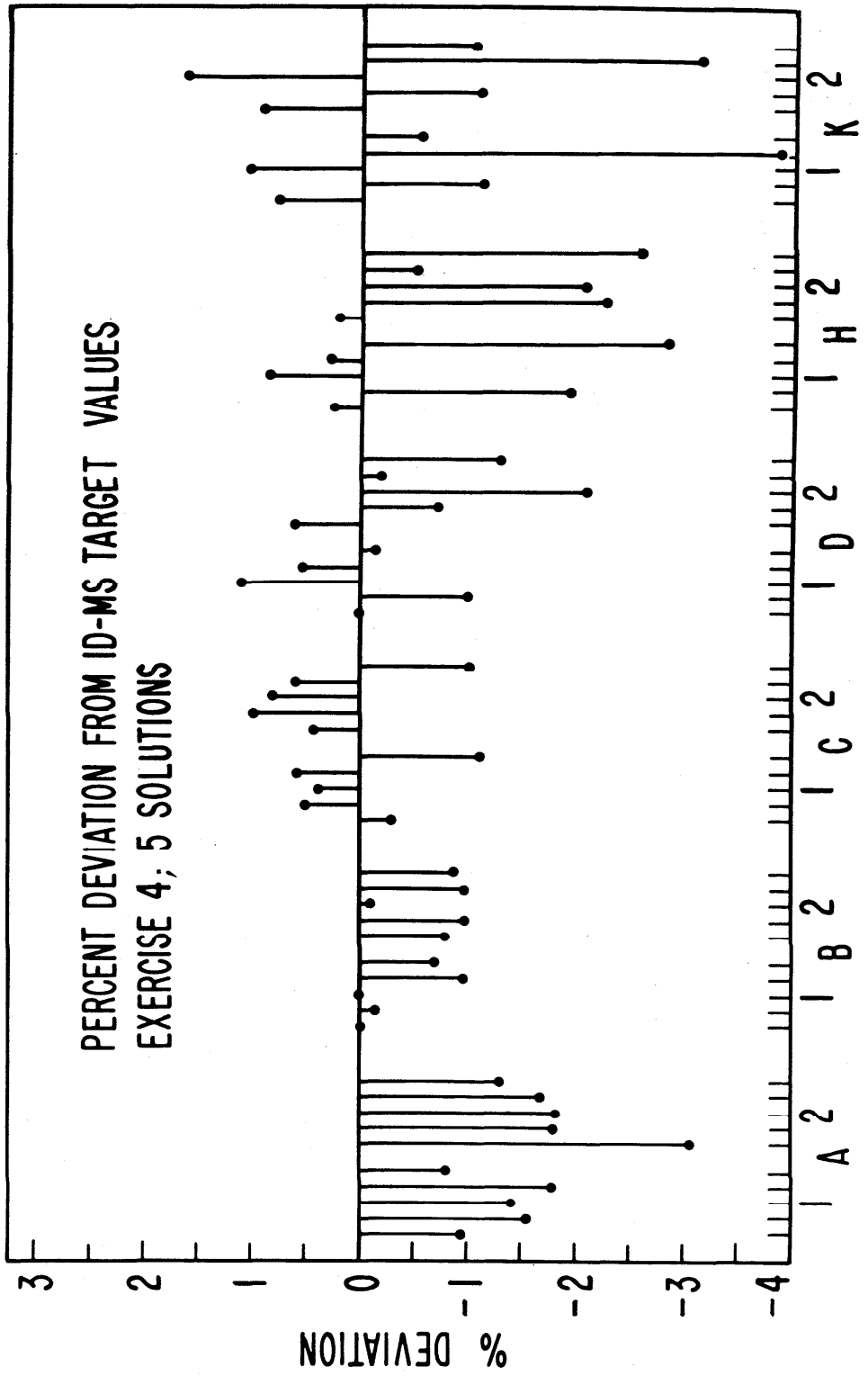

Figure 5. Laboratories Code and Round Numbers

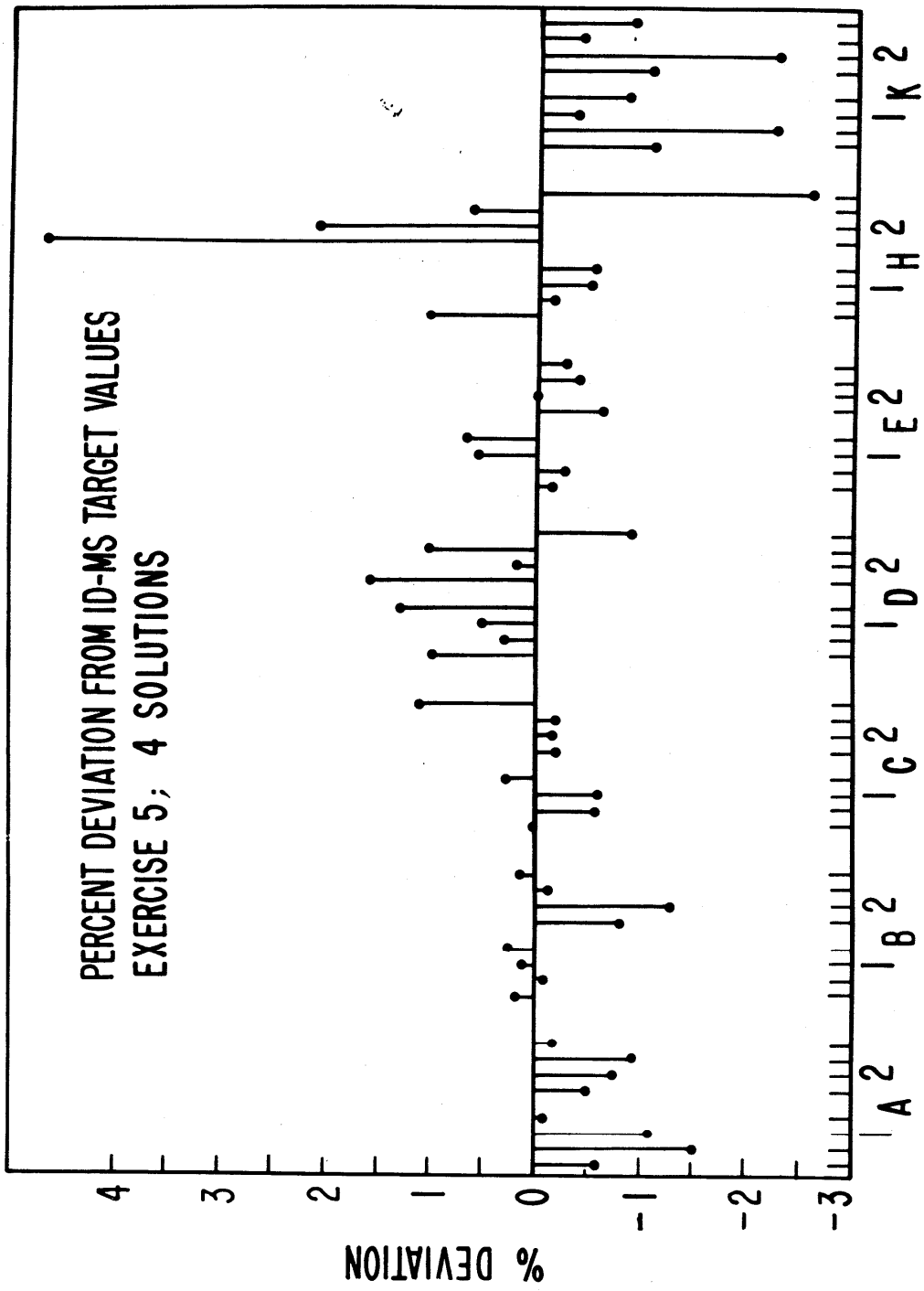

Figure 6. Laboratories Code and Round Numbers

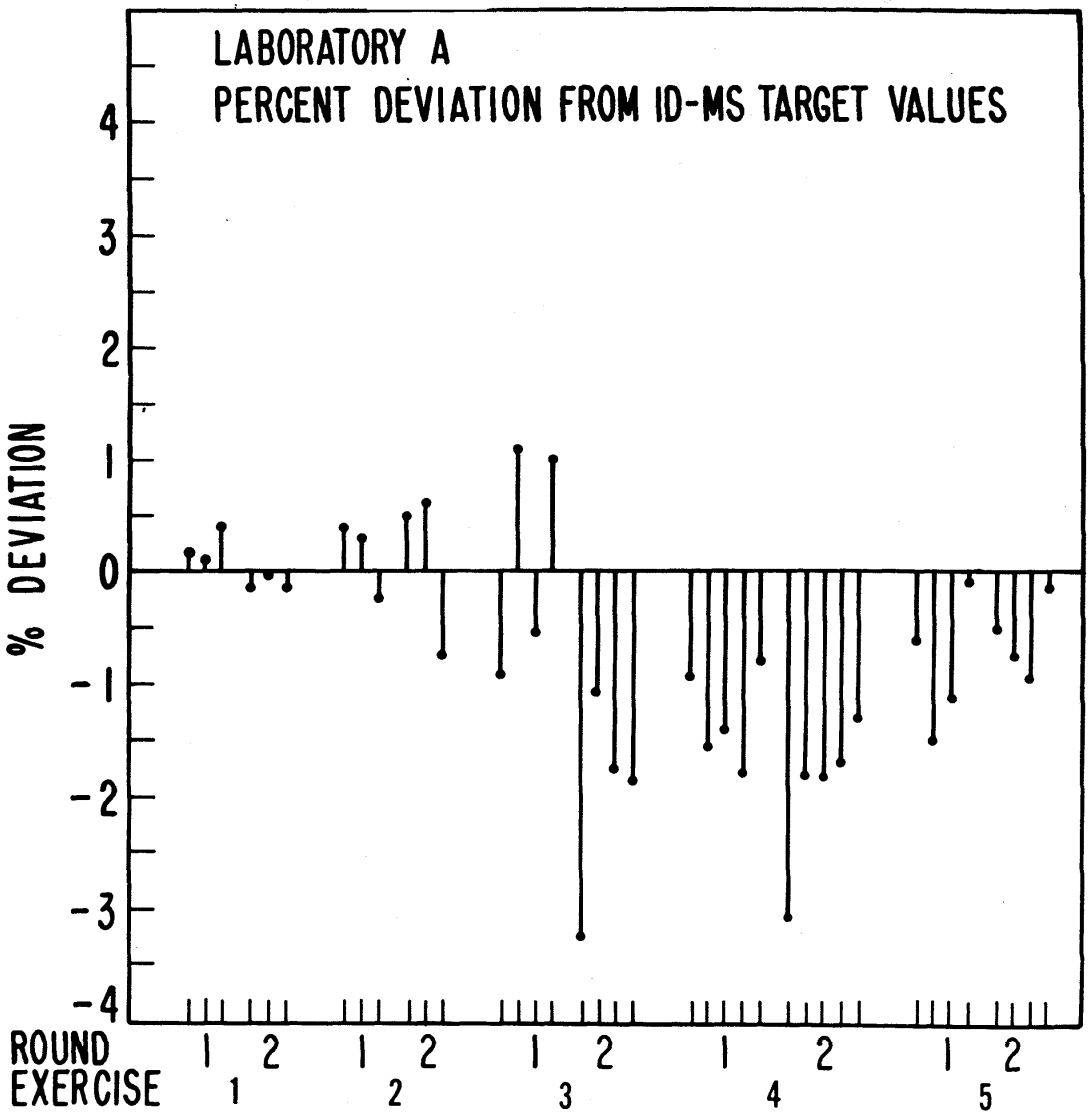
LABORATORY A
PERCENT DEVIATION FROM ID-MS TARGET VALUES
% DEVIATION
4
3
2
1
0
-1
-2
-3
-4
ROUND
EXERCISE
1 2 1 2 1 2 1 2 1 2
1
2
3
4
5

Figure 7.

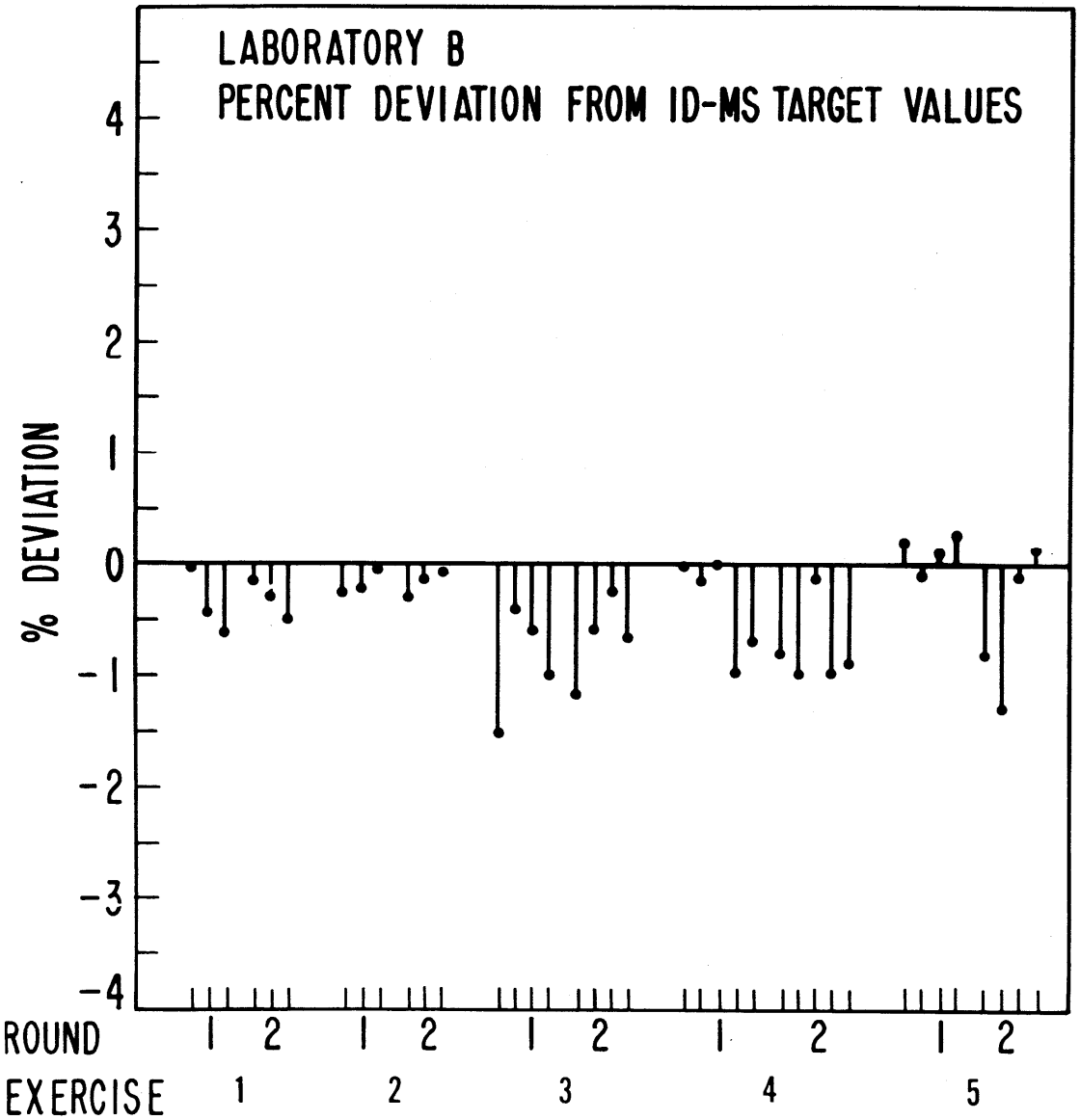
LABORATORY B
PERCENT DEVIATION FROM ID-MS TARGET VALUES
% DEVIATION
4
3
2
1
0
-1
-2
-3
-4
ROUND
1 2 1 2 1 2 1 2 1 2
EXERCISE
1 2 3 4 5

Figure 8.

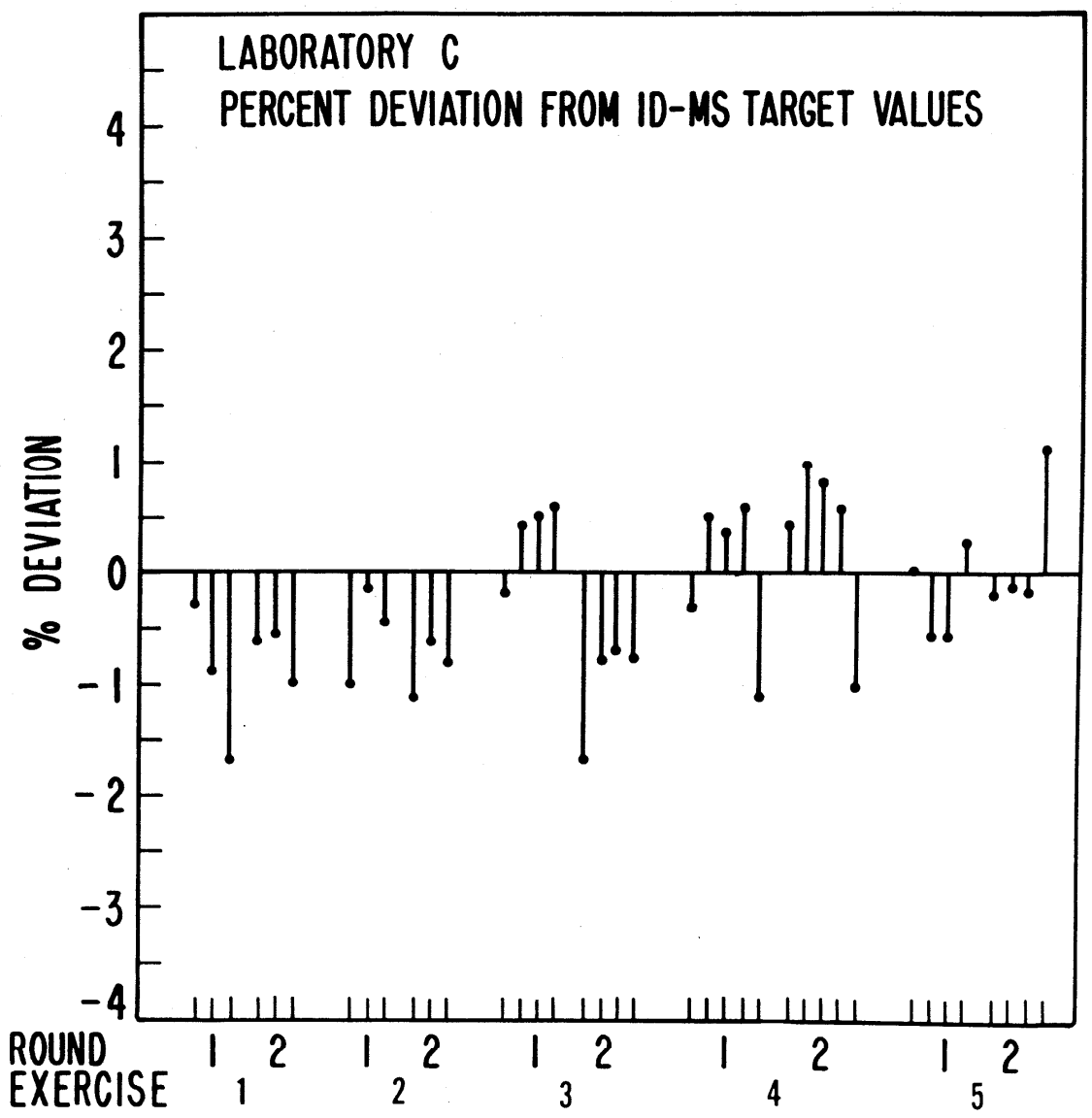
LABORATORY C
PERCENT DEVIATION FROM ID-MS TARGET VALUES
% DEVIATION
4
3
2
1
0
-1
-2
-3
-4
ROUND
EXERCISE
1 2 1 2 1 2 1 2 1 2
1
2
3
4
5

Figure 9.

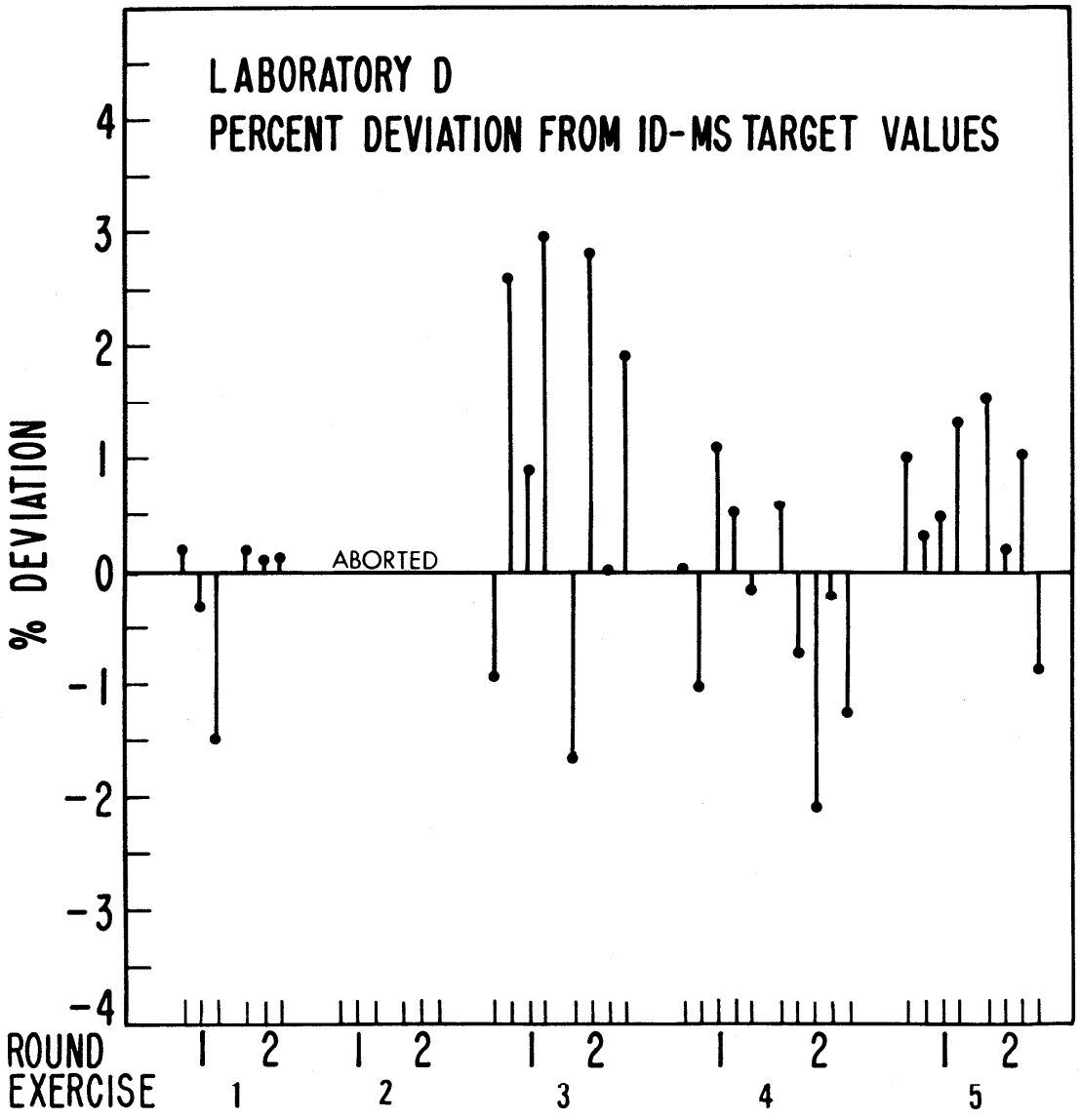
LABORATORY D
PERCENT DEVIATION FROM ID-MS TARGET VALUES
% DEVIATION
4
3
2
1
0
-1
-2
-3
-4
ABORTED
ROUND
EXERCISE
1 2 1 2 1 2 1 2 1 2
1
2
3
4
5

Figure 10.

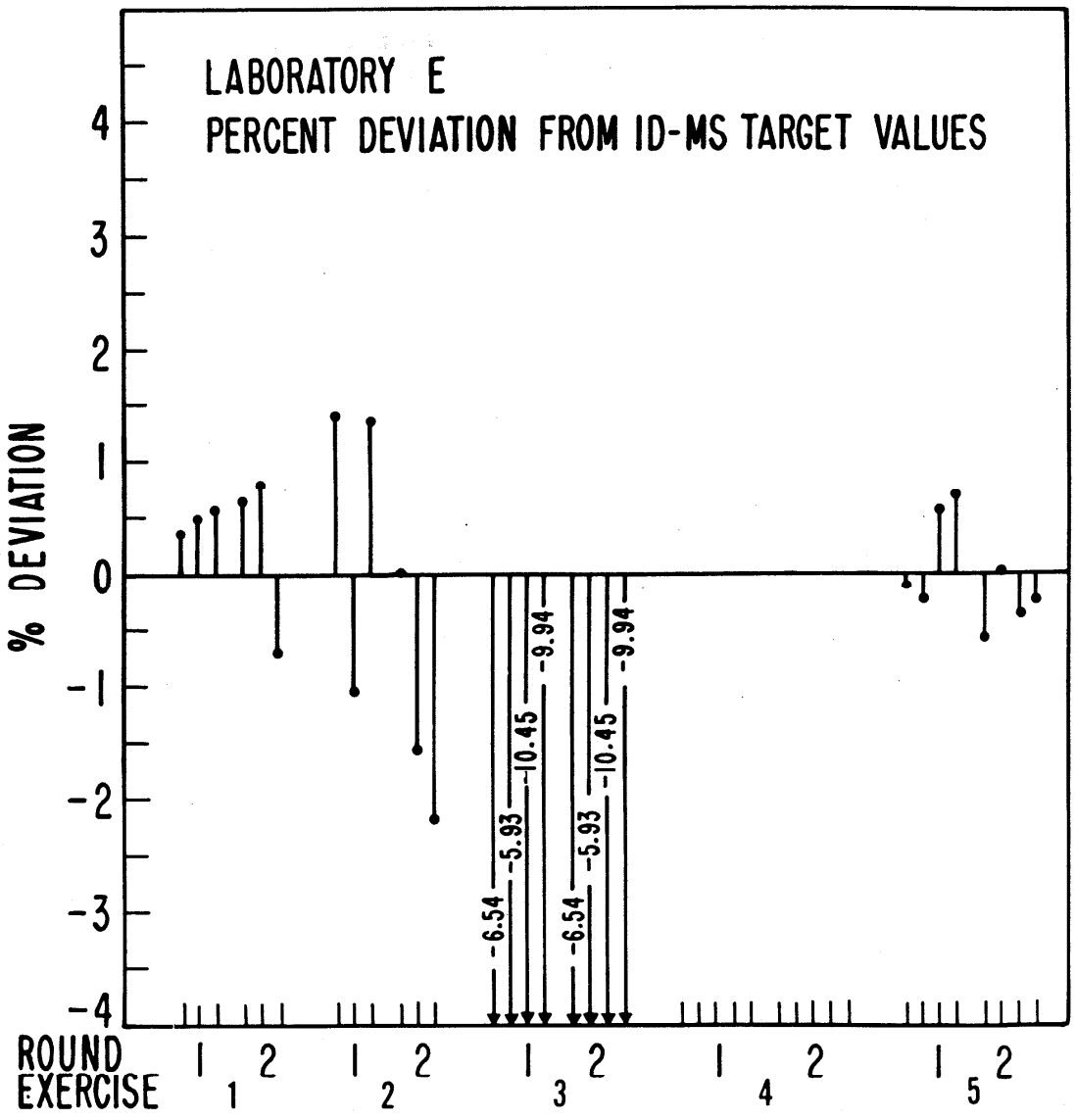
LABORATORY E
PERCENT DEVIATION FROM ID-MS TARGET VALUES
% DEVIATION
4
3
2
1
0
-1
-2
-3
-4
-6.54
-5.93
-10.45
-9.94
-6.54
-5.93
-10.45
-9.94
ROUND
EXERCISE
1 2
1 2
1 2
1 2
1 2
1
2
3
4
5

Figure 11.

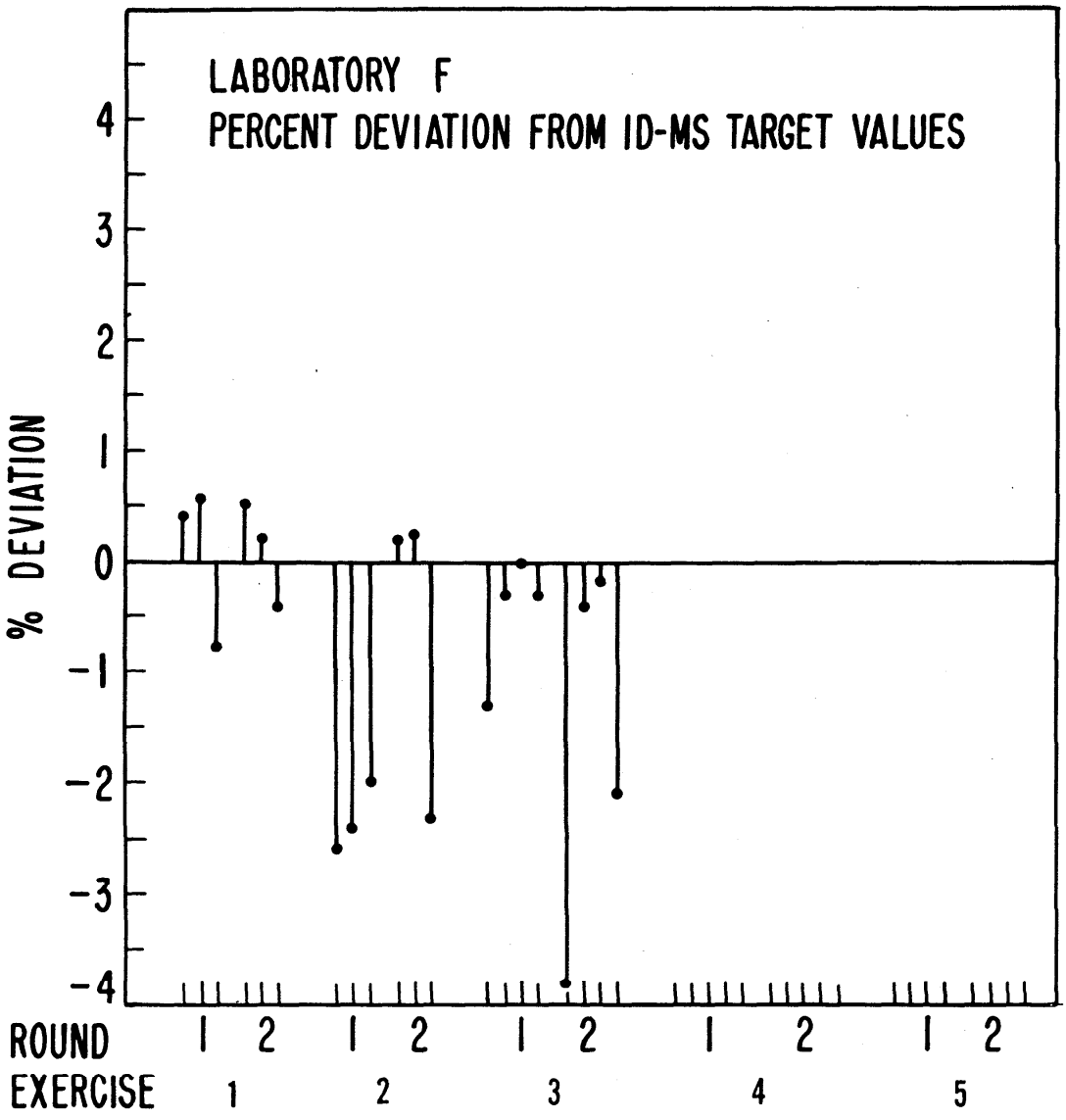
LABORATORY F
PERCENT DEVIATION FROM ID-MS TARGET VALUES
% DEVIATION
4
3
2
1
0
-1
-2
-3
-4
ROUND
EXERCISE
1 2 1 2 1 2 1 2 1 2
1
2
3
4
5

Figure 12.

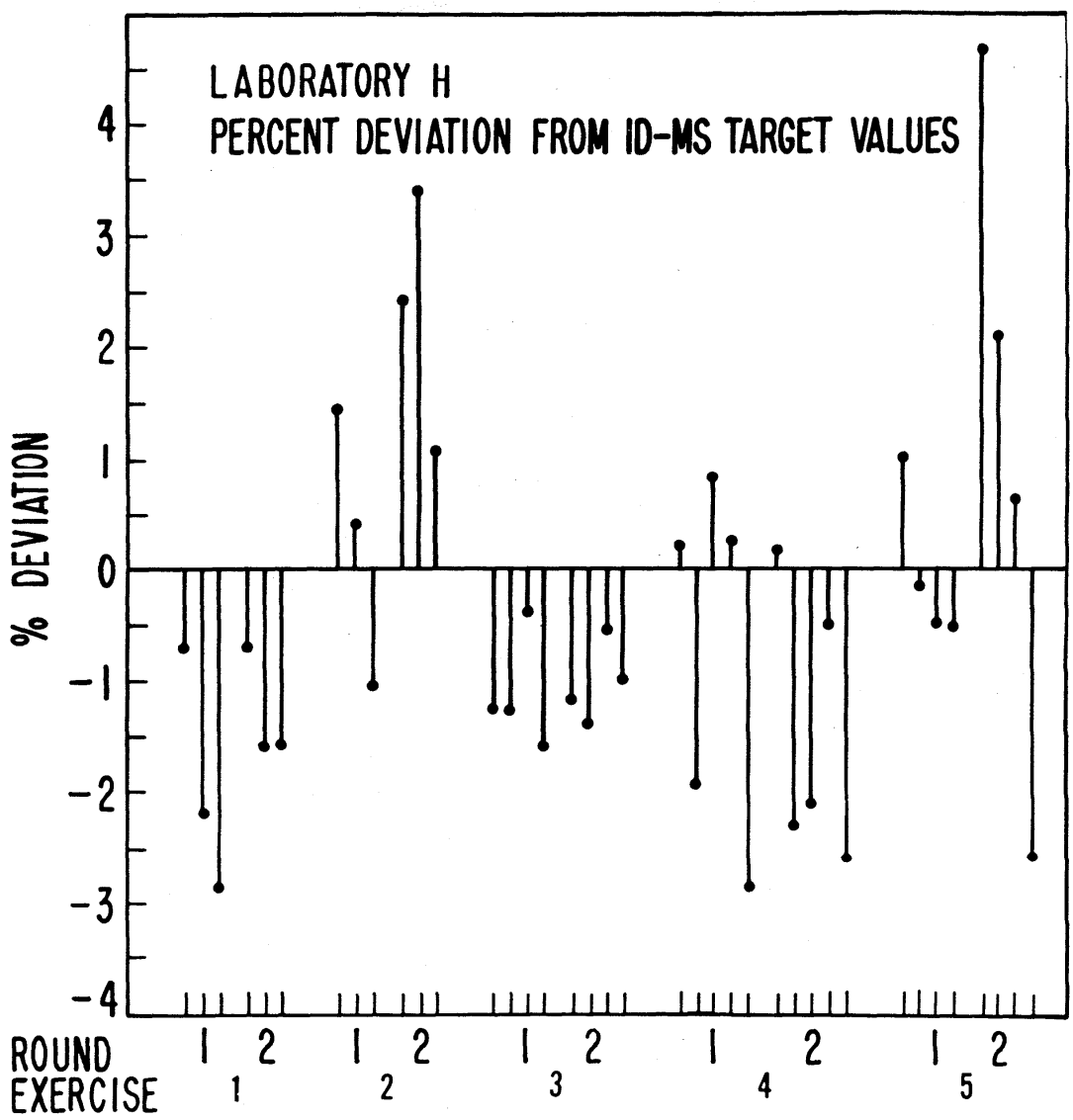
LABORATORY H
PERCENT DEVIATION FROM ID-MS TARGET VALUES
% DEVIATION
4
3
2
1
0
-1
-2
-3
-4
ROUND
EXERCISE
1 2 1 2 1 2 1 2 1 2
1
2
3
4
5

Figure 13.

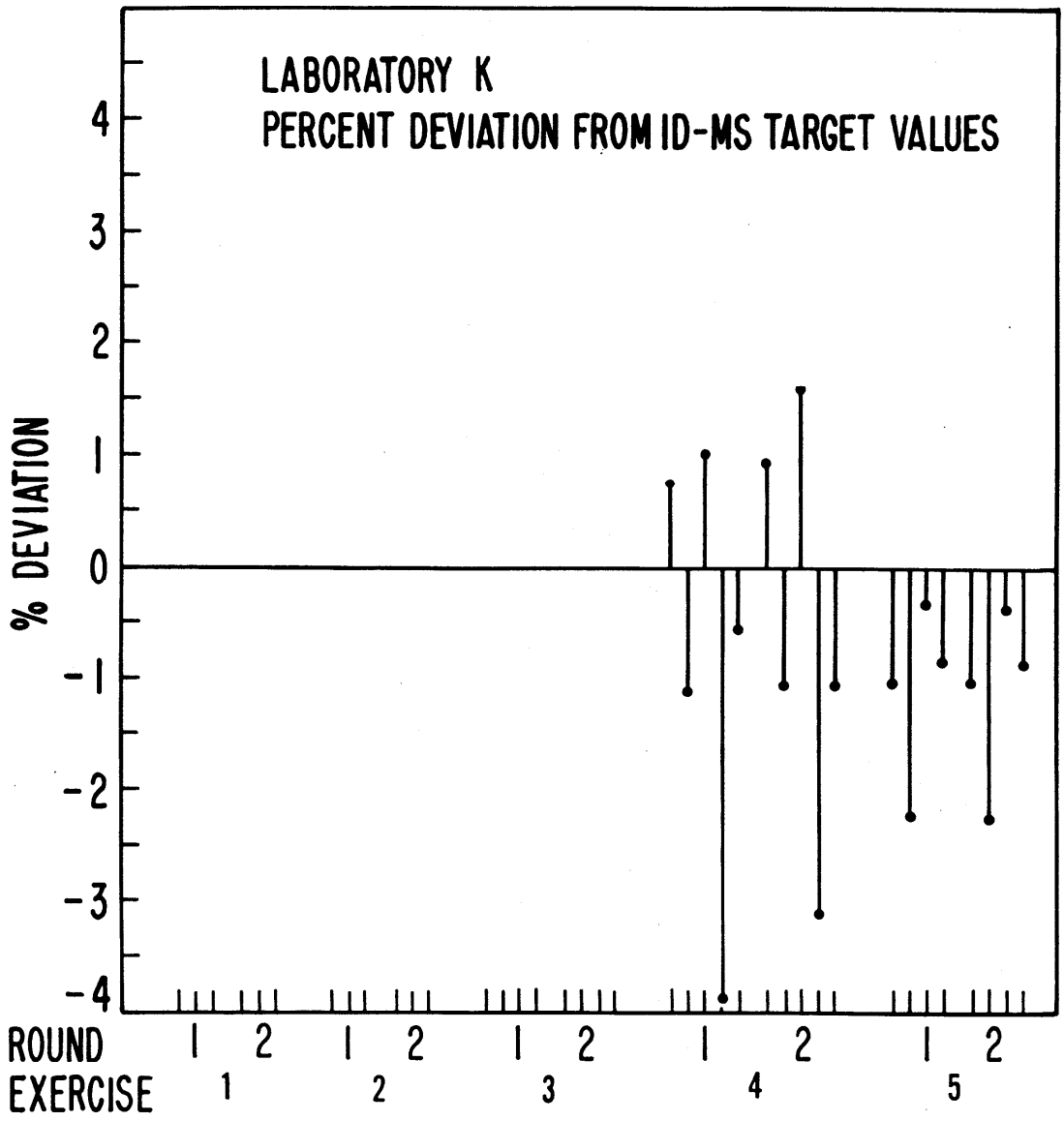
LABORATORY K
PERCENT DEVIATION FROM ID-MS TARGET VALUES
% DEVIATION
4
3
2
1
0
-1
-2
-3
-4
ROUND
EXERCISE
1 2 1 2 1 2 1 2 1 2
1
2
3
4
5

Figure 14.

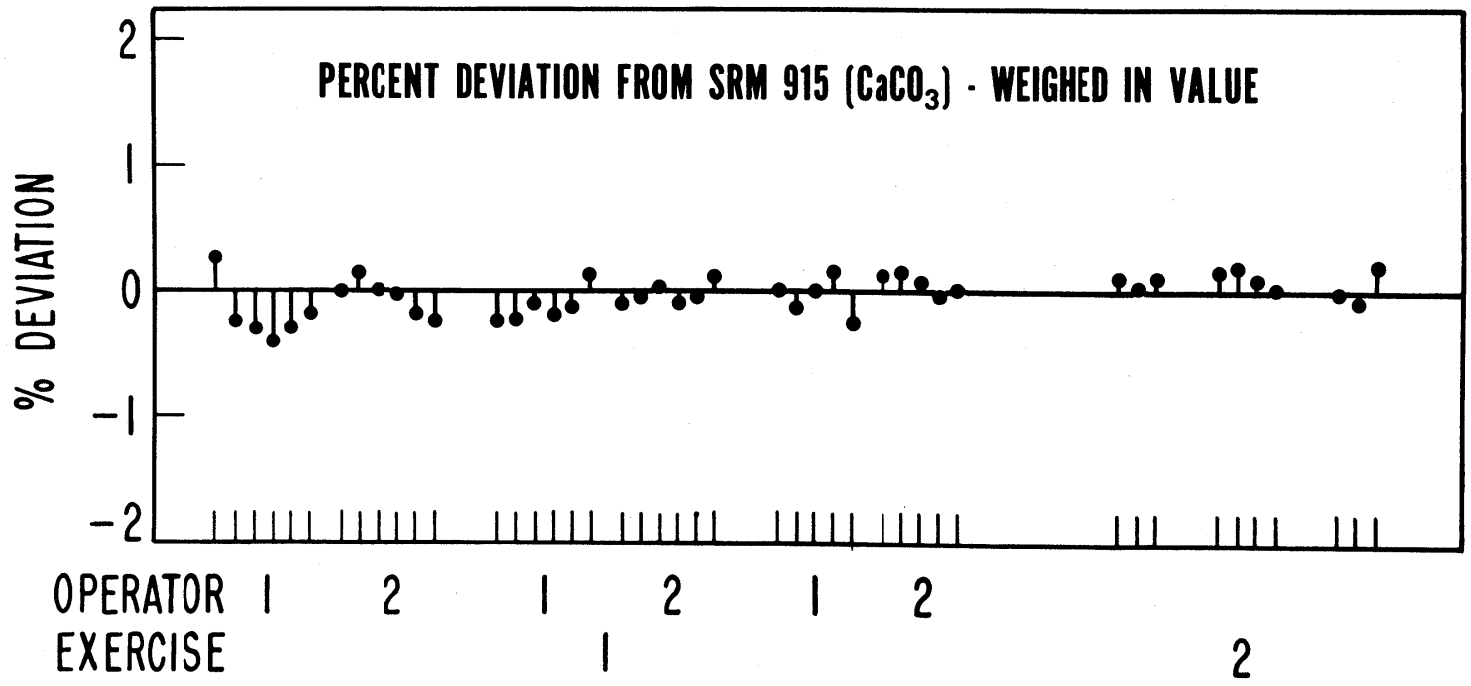
PERCENT DEVIATION FROM SRM 915 ($CaCO_3$) - WEIGHED IN VALUE
% DEVIATION
2
1
0
-1
-2
OPERATOR 1 2 1 2 1 2
EXERCISE 1 2

Figure 15.

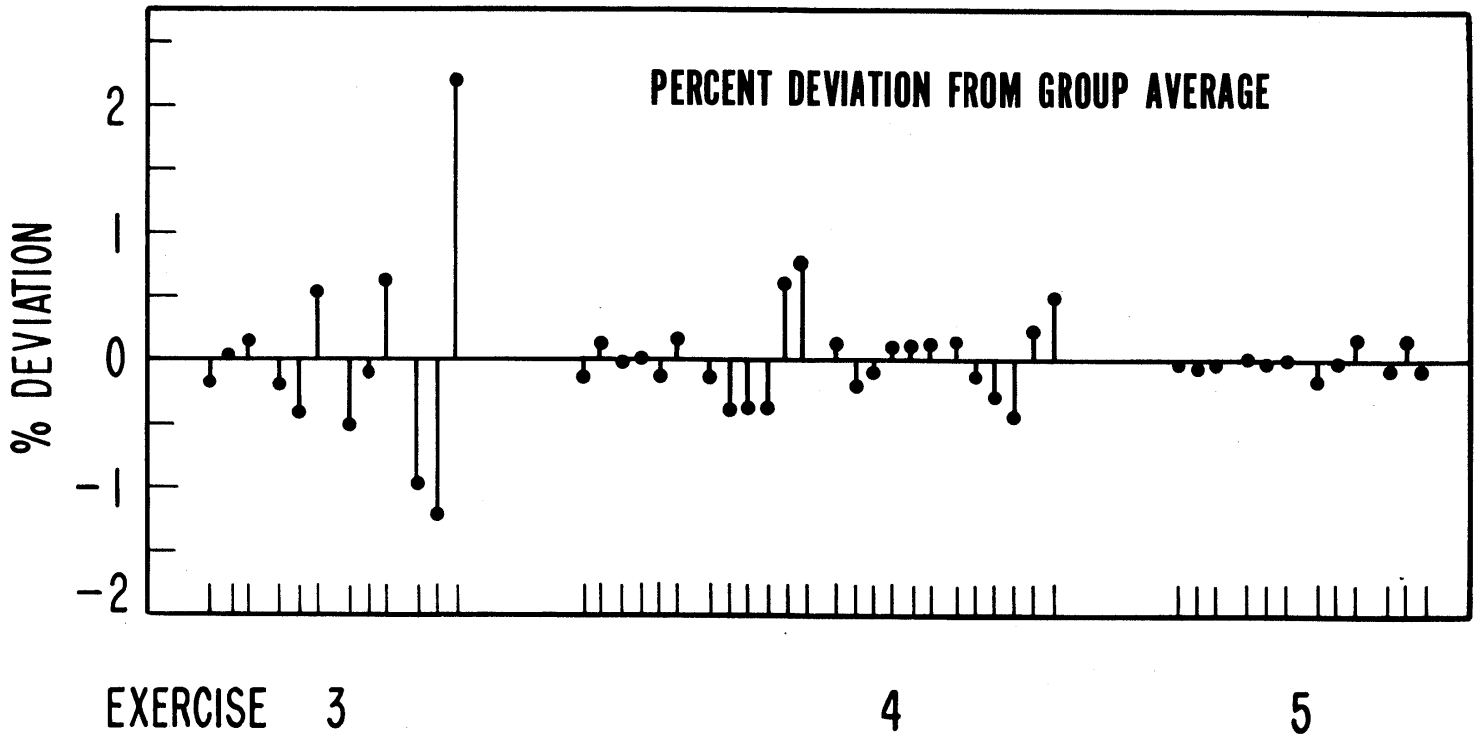
PERCENT DEVIATION FROM GROUP AVERAGE
% DEVIATION
2
1
0
−1
−2
EXERCISE
3
4
5

Figure 16.

VII. THE ISOTOPE-DILUTION MASS SPECTROMETRY WORK

(Larry J. MOORE)

A. Theory of Stable Isotope Dilution Analysis

The analysis of trace level components by stable isotope dilution mass spectrometry is a well-established technique and has been successfully applied to a wide variety of materials over the last three decades. Any analytical method capable of measuring the relative quantity of stable isotopes in an element is a potential candidate for isotope dilution analyses; however, mass spectrometry has been used almost exclusively due to its sensitivity and accuracy. More specifically the thermal ionization technique has been the predominant mass spectrometric method used because of its inherent accuracy, sensitivity and versatility for the measurement of inorganic elements.

Isotope dilution analyses are performed by measuring the change in the relative magnitude of two analyte isotopes when a measured amount of a material enriched in one isotope is added to the sample. The method of calculating the concentration of the analyte from an isotope pair ratio measurement is shown in equation 1.

$$\text{Concentration, ppm (wt)} = \frac{W_{sp}\, C[A_{sp} - R\, B_{sp}]}{BR - A} \cdot \frac{M}{W_s} \qquad (1)$$

W_{sp} = Weight of spike solution, grams
C = Concentration of spike, μmoles/gram of solution
A_{sp} = Atomic fraction of isotope A in spike
B_{sp} = Atomic fraction of isotope B in spike
A = Atomic fraction of isotope A in sample
B = Atomic fraction of isotope B in sample
R = Experimentally measured ratio
M = Atomic weight of analyte
W_s = Weight of sample, grams

Upon dissection of this equation it is observed that a successful and accurate isotope dilution concentration measurement must include the following steps:

(1) Addition of the spike to the analyte, chemical dissolution and (usually) treatment of the sample to eliminate potential mass spectrometric interferences.

(2) Accurate measurement by mass spectrometry of the isotopic composition of both the analyte as it exists naturally in the sample and the enriched isotope mixture used for altering the natural isotope ratio.

(3) Determination of the concentration of the spiking isotope solution by calibrating it versus accurately known amounts of a solution of the "natural" analyte.

(4) Determination of the amount of natural analyte contamination (blank) incurred during the chemical processing.

(5) Measurement of the altered isotope ratio and calculation of the concentration.

B. Mass Spectrometric Isotope Dilution as an Absolute Analytical Technique

1. Systematic Errors Possible in Isotope Dilution Measurements

a. Chemical contributions

A critical evaulation of the errors possible in the various chemical treatment steps results in several contingencies. The two most frequent forms of systematic error in chemical processing are (1) incomplete dissolution of sample and (2) the chemical blank or contamination incurred during the processing.

For concentrations where the chemical blank represents a significant fraction of the total analyte, the blank may

become a contributing error. In fact, when all the other contributing systematic and random errors of a total isotope dilution method are under control the blank may become the limiting error.

The blank can be dissected as consisting of contributions from reagents and particulate fallout. To reduce contributions from the first source, ultra-pure acids (Ca present in $HClO_4$, HNO_3 and HCl at 0.2 ppb-ng/g-0.2 ppb and 0.06 ppb, respectively) are prepared and used at NBS. To reduce particulate contributions all the chemical preparations are carried out in a Class 100 clean air hood inside a vertical flow clean room [7].

If repetitive blank determinations for a method are uncertain to $\leq\sim$50% the blank can be treated as a systematic error whose uncertainty as a correction is equal to the randomness of its measurement. Or, if the blank is uncertain to $\geq\sim$100% then the total blank may become the limit of error for trace concentrations and in that case must be applied to the concentration result as a plus or minus limit of error. Of course, the error from a given blank level becomes larger as the analyte concentration becomes lower. The contribution of the blank correction to the total error therefore depends on (1) the concentration of the analyte, (2) the reproducibility of the blank, and (3) the magnitude of competing error contributions from mass spectrometric and other chemical sources, as described in this section. The importance of recognizing the random blank as a basic limit of error in many trace element analyses, regardless of the analytical technique, cannot be overemphasized. Its identification, control and correction should be considered an integral part of the analysis.

In addition to incomplete dissolution of the sample, there are other sample-oriented errors that can also occur if reasonable care is not taken by the chemist in sample preparation:

(1) isotope fractionation occurring during non-quantitative sample recovery, (2) incomplete mixing of the sample with the separated isotope, and (3) partial loss of the analyte upon sample dissolution.

As seen in equation 1, the accuracy of the determination is directly dependent on the accuracy of the spike calibration. As the calibration uses a known "sample", most of the sample-oriented errors apply here as well.

If the possible errors outlined above are considered in the careful design and application of an experiment, they can be eliminated as sources of systematic bias.

b. Mass spectrometric contributions

The formation of ions in a triple filament thermal ionization source is accomplished by heating the sample on each of two 1x30 mil rhenium ribbons to volatilize the sample to the relatively hot surface of a third rhenium ribbon where a portion of the atoms leave the surface as ions. The ionization process is a very gentle one, resulting in a distribution of excess ion energies of only a fraction of an electron volt, which is an order of magnitude less than its closest competitor, the electron impact source. The low energy spread of the ions and the stability of the ion beam are the bases for the inherent capability of the thermal ionization technique to make accurate isotope measurements.

Errors in the measurement of isotope ratios can occur from several sources. The ability of the operator to control fractionation of the analyte isotopes in the ion source reproducibly from one sample to the next is usually the limiting error in isotope dilution analyses. In the case of $^{40}Ca/^{44}Ca$ measurements, the difficulty in controlling isotope fractionation is compounded by a basic limitation in the measurement of isotopes involving large mass differences (difference of 10% for ^{40}Ca and ^{44}Ca).

Instrumental errors can occur in the non-linear measurement of varying isotope ratios. However, the instruments used in this work have been repeatedly shown to be bias-free in the measurement of ratios varying from 1/20 to 20/1 at a 95% limit of error (calculated as $t\sigma$) of ±0.05% [8]. Thus any contribution from this source is negligible since the $^{40}Ca/^{44}Ca$ reproduces only to ∿±0.2% (95% L.E.), which is a factor of four higher than ±0.05%.

Background contributions from the filament material at the mass positions of interest (in the present case, 40 and 44) can be a source of error and must be investigated.

Interferences from isobars at the mass positions of interest can be a problem (e.g. ^{40}K is isobaric with ^{40}Ca) and must be controlled. Usually the sample is processed chemically to rid it of isobaric interferences, although occasionally the interference can be eliminated during mass spectrometric analysis.

Particulate contamination can occur during loading of the sample prior to mass spectrometric analysis. Sample loading is therefore carried out insofar as possible inside a Class 100 clean air hood and care is taken to avoid excess handling of sample-holding equipment.

c. General considerations

When the potential sources of systematic bias are eliminated, the residual error is composed of the random errors associated with the measurement of the isotope ratios and with the chemical preparation steps. If the blank correction is sufficiently large and random, then it must compete with the total random error derived from the other sources to become the limiting error for the analysis. With these qualifications, mass spectrometric isotope dilution analysis then becomes a method for the absolute analysis of trace elements in a variety of matrices.

2. Historal Performance of Isotope Measurements

Because of its inherent accuracy, thermal ionization mass spectrometry has found extensive application at NBS for the measurement of isotope ratios. In the process of making these measurements the Analytical Mass Spectrometry Section has invested, conservatively, 20-30 man-years in the elucidation of the various error components. The end result of these efforts has been the absolute determination of the atomic weights of Mg, Pb, Cr, Cl, Ag, Rb, B, Cu and Br [9-17]. In addition, K, Ni, Sr, Re and Si are currently under development for atomic weight determinations.

Over the last few years the knowledge acquired during these isotope ratio measurements has been applied widely to isotope dilution analyses. As an example, the certification of the Trace Elements in Glasses standards has required the development of isotope dilution analysis methods for U, Th, Pb, Tl, Rb and Sr among others at concentration levels of 500-0.02 ppm. Accuracies for some elements in these standards have been maintained at $\leq\sim$0.25% (95% L.E.), even at concentrations less than 1 ppm (publications in preparation). Applications of isotope dilution have also been made to a wide variety of other matrices, from beef liver and orchard leaves to nuclear fuel rods and reference steels. In all these applications the rules by which isotope measurements are made have not been found to change, and are directly applicable to the determination of calcium in serum.

C. Analytical Procedure

1. Chemistry

The individual samples for analysis by atomic absorption in Exercise 1 were taken from stock solutions of SRM 915 $CaCO_3$ which were prepared at concentrations of approximately 80, 100 and 120 ppm calcium. Aliquoting of the samples was accomplished by transferring 50 ml portions of the stock solution to 100 ml borosilicate glass ampoules followed by sealing of the ampoules. No error from weight

loss was observed due to evaporation of the solution during the sealing process.

Ampoules from the beginning, middle and end of the aliquoting procedure were taken for isotope dilution analysis and their pH and density were measured. Samples were obtained from the ampoules by multiply extracting 5 gram quantities with a syringe, which were then accurately weighed. These samples were spiked with a known amount of the ^{44}Ca separated isotope and converted to the nitrate form for isotopic analysis.

Solutions for Exercise 2 were prepared and ampouled as in the first exercise, except that sodium and potassium were added prior to ampouling at 140 mg/l and 5 mg/l, respectively, to simulate the ionic strength of serum. The addition of the alkali elements required that the samples taken for isotope dilution be processed by ion exchange to remove the potassium, since ^{40}K is isobaric with ^{40}Ca. Otherwise the isotope dilution samples were processed as in Exercise 1.

To obtain atomic absorption samples for Exercise 3, 50 ml aliquots were transferred to 100 ml ampoules from four 1-liter lots of pooled blood serum supplied by Hartford Hospital. Multiple samples for isotope dilution were withdrawn by syringe from selected ampoules and $HClO_4$ was added to destroy the organic matter present in the serum. The calcium was purified by ion exchange and the isotope ratios determined by mass spectrometry.

Exercise 4 was carried out as in the third exercise, again with pooled serum samples supplied by Hartford Hospital. No significant changes in the chemical procedure were required, except that a larger bore syringe was used to avoid sampling problems described later.

The aliquoting, sampling and analysis in Exercise 5 were performed as before, but with samples supplied by the Center

for Disease Control.

Insofar as possible, all of the above chemical operations were carried out in a Class 100 clean air hood. Following the construction of a vertical flow clean room midway through the series of analyses all chemical operations were transferred to this facility.

2. Mass Spectrometry

All calcium isotope ratio measurements were made using three magnetic sector, single focussing thermal ionization mass spectrometers: one was a 12" radius of curvature, 68° sector instrument, and the remaining two were identical 6" radius of curvature, 68° sector instruments. All the instruments use identical ion sources and deep faraday cup collectors [18-19].

Care was taken to maintain an identical heating and fractionating pattern for the analysis of both standards and samples to ensure no systematic bias from this source.

Filament material background contributions from ^{40}Ca and interferences from ^{40}K were carefully investigated and removed.

Calcium samples and standards (SRM 915) were loaded as calcium nitrate onto each of two rhenium sample filaments. Using a triple filament rhenium technique (the details of the chemical and mass spectrometric techniques used in the analysis of Ca by isotope dilution are being published separately), the samples were analyzed for their $^{40}Ca/^{44}Ca$ ratios, and the concentrations calculated according to Equation 1.

D. Analysis Results

The following compendium of data is composed of the isotope dilution concentrations for Exercises 1-5, expressed as milliequivalents of Ca per liter. As all of the isotope dilution calculations in the Analytical Mass Spectrometry Section are performed on a weight basis; i.e., ppm (wt), the density was carefully measured for each concentration level

in each exercise to permit a conversion of the ppm concentrations to milliequivalents/liter:

$$\text{Concentration, milliequivalents/liter} = \frac{(\text{ppm})\ (D)}{EW}$$

ppm = Calculated concentration from Equation 1, μg/g
D = Measured density, g/ml
EW = Equivalent weight of calcium (20.04).

Exercise 1 (Table 6)

Exercise 1 was the first attempt of the Analytical Mass Spectrometry Section at the isotope dilution analysis of calcium. Two mass spectrometers were used, each operated by different analysts using the same mass spectrometric technique. $^{40}Ca/^{44}Ca$ isotope ratio representing the average of 10 ratio measurements was determined for each sample, and the average ratios that produced outlying concentration values were repeated. Because these solutions were prepared using primary chemical methods, the agreement among the aliquots for a given concentration level represents a direct measure of the reproducibility of the isotope dilution technique.

Using 16 degrees of freedom, a 95% L.E. was calculated for each operator's concentration values*. The magnitude and variation of the operators' measurement uncertainties (0.35% and 0.22%) are due to the fact that the analysis and the development of an analytical technique for calcium were carried out concurrently. Normally a technique is investigated and developed over a period of several months, and only then is the perfected technique applied to accurate isotope measurements.

---

*The ratio of individual concentration-to-average concentration was computed for each sample for each concentration level and the variations from 1.0000 were used for calculating the 95% L.E. for a single analysis.

From previous experience, it has been found that two operators are often useful in establishing the absence of any systematic bias introduced by the operator. Although there was no detectable operator bias indicated in the measurements, the average concentration values of each operator were averaged to obtain the final reported values. The agreement between the experimental and calculated values may be taken as an experimental verification that the possible systematic errors have been eliminated and that the calculated confidence limits are a reflection of the random errors associated with the isotope ratio measurements.

The remaining possible systematic error is that of the blank, which was found to be reproducible at $\leq\sim 0.1\%$ of the total amount of calcium per sample. Again, the excellent agreement between calculated and experimental values indicates that the blank is insignificant as a source of systematic bias.

Each sample number in Exercise 1 (1, 1A, etc.) represents a separate chemical preparation. In this exercise three ampoules from each concentration level (e.g. vial numbers 1, 20, 120) were sampled in duplicate (1, 1A, 20, 20A, etc.). Although not all the prepared samples in later exercises were analyzed for isotope ratios, the same sampling nomenclature is followed throughout Exercises 1-5.

Table 6. Concentration of Calcium in Synthetic $CaCO_3$ Solutions, meq/liter.

Exercise 1.

Sample	Operator 1	Operator 2	Average	Calculated	Relative error
1	4.093	4.083	4.088		
1A	4.072	4.089[a]	4.081		
20	4.071	4.083	4.077		
20A	4.066	4.081	4.074		
120	4.071	4.075	4.073		
120A	4.075	4.073	4.074		
Average	4.075	4.081	4.078	4.083	0.12[b]
21	5.173	5.180	5.177		
21A	5.174	5.183	5.178		
61	5.180	5.187	5.183		
61A	5.176	5.180	5.178		
80	5.179	5.183	5.181		
80A	5.192[a]	5.192	5.192		
Average	5.179	5.184	5.182	5.186	0.08
41	5.867	5.874	5.871		
60	5.859	5.876	5.868		
60A	5.867	5.870	5.869		
99	5.875	5.865	5.870		
99A	5.852	5.867	5.859		
Average	5.864	5.870	5.867	5.867	0.00
95% L.E.	±0.35%[c]	±0.22%[c]			

[a] Isotope ratios determined in duplicate.

[b] Relative error is the difference between experimental and calculated value expressed as a percentage of the calculated value.

[c] Error expressed as 95% L.E. for a single analysis.

Exercise 2 (Table 7)

These results were obtained using a single mass spectrometer, and involved the measurement of a smaller number of aliquots for each concentration level. A 95% L.L. was calculated for the results in Exercise 2 in a manner analogous to that of Exercise 1. Even though fewer degrees of freedom were used the uncertainty for the measured concentrations was less than for Exercise 1, as a direct result of increased experience in the isotope measurements.

Also of significance in this exercise, the extra chemical steps involved in the separation of calcium from sodium and potassium resulted in no measurable introduction of a systematic bias. A proof of the absence of any new systematic bias is reflected in the agreement ($\leq\sim 0.1\%$) between calculated and experimental concentrations. The blank correction for this series was $\leq\sim 0.1\%$ of the total amount of calcium per sample.

In summary, Exercises 1 and 2 provided a testing ground for experimentally demonstrating the validity of isotope dilution as a technique for the absolute analysis of calcium at approximate serum concentrations. By carefully controlling and eliminating the well-known bias factors, the concentration of calcium was determined with a consistent accuracy of $\leq\sim 0.1\%$.

Exercise 3 (Table 8)

A direct and successful extrapolation of the isotope dilution technique to serum samples at the desired accuracy level of $<0.5\%$ was precluded in this exercise by poor samples. Upon thawing, the several lots of serum succumbed to bacteriological action that resulted in turbid serum samples with differing amounts of suspended particulate matter of varying size. As a consequence, sampling by weight with a relatively narrow bore syringe was an apparent source of sampling error. The syringe had the effect of filtering out all particles too large to be admitted into the syringe needle.

Table 7. Concentration of Calcium in Synthetic $CaCO_3$ Solutions (Na and K added), meq/liter Exercise 2.

Sample	Isotope dilution	Calculated	Relative error
231	3.593		
235	3.591		
276	3.594		
Average	3.593	3.590	0.08[a]
215	4.895		
215A	4.896		
242	4.891		
242A	4.887		
Average	4.892	4.887	0.10
203A	6.374		
263	6.369		
266	6.388		
Average	6.377	6.375	0.03
95% L.E.	±0.20%[b]		

[a] Relative error is the difference between experimental and calculated value expressed as a percentage of the calculated value.

[b] Error expressed as 95% L.E. for a single analysis.

The ability to measure calcium concentrations by isotope dilution with a 95% L.E. per analysis of ∿0.2% can be inferred from Exercise 2. In addition, the uncertainty of the concentration values of this exercise was diminished by making a minimum of duplicate isotope measurements per sample. Therefore, any significant excursions from the 0.2% accuracy level were largely attributable to sampling. Although not all of the data obtained in Exercise 3 are shown here, it was readily apparent that serious aliquoting errors occurred at the 1-3% level. The largest error of ∿3% was in the 308-316 samples, while the remaining levels exhibited ranges $\leq$∿1%.

Since the dominant error on this exercise was due to poor sampling, which in turn varied among the lots of serum, the error can only be estimated as plus or minus the range of aliquots for each concentration level.

Exercise 4 (Table 9)

To circumvent the sampling problems encountered in the previous exercise a large bore syringe was used for sampling and the serum was kept frozen until just prior to use. In general, the results were much more encouraging. Although apparent sampling problems still existed in aliquots 561-576 and 501-516, the remaining aliquot groups (521-535 and 541-554) ranged <∿0.3%. An aliquoting error is easily seen in 516C and 516D, which are duplicate subsamples of one aliquot for the 501-516 group. Aliquots 501 (C and D) and 508 (C and D) agree internally and with each other to <∿0.25%, but differ collectively from 516 (C and D) by ∿1%. The blank correction was <0.1% for all samples.

Again, due to poor sampling, the error could only be estimated as plus or minus the range of aliquot values per concentration level.

Table 8 Concentration of Calcium in Serum, meq/liter

(Exercise 3.)

Sample		Isotope dilution
301		6.094[a]
308		6.079
316		6.288[b]
	Average	6.154 ±3.40%[c,d]
321		4.560
328		4.570
333		4.575
	Average	4.568 ±0.33%[c]
341		4.999
348		5.020
356		5.056
	Average	5.025 ±1.11%[c]
361		4.818
369		4.808
377		4.859
	Average	4.828 ±1.06%[c]

[a] All isotope ratios determined in duplicate (except 316).

[b] Isotope ratio determined in quadruplicate.

[c] Errors estimated as plus or minus the range of individual aliquots for each concentration level.

[d] Sampling is the dominant error for the results in Exercise 3. The 95% L.E. for the isotope dilution analysis of calcium is ±0.20%, as shown in Exercise 2.

Table 9. Concentration of Calcium in Serum meq/liter.
(Exercise 4)

Sample	Isotope dilution	Sample	Isotope dilution
521-C	5.065	541-C	3.616
521-D	5.050[a]	541-D	3.626
528-C	5.055	548-C	3.621
528-D	5.065	548-D	3.621
535-C	5.065	554-C	3.616
535-D	5.065	554-D	3.626
Average	5.061 ±0.30%[b,c]		3.621 ±0.28%[b,c]
561-C	5.801	501-C	4.106
561-D	5.785	501-D	4.096
568-C	5.775	508-C	4.096
568-D	5.765[a]	508-D	4.096[a]
576-C	5.806	516-C	4.136
576-D	5.821	516-D	4.142[a]
Average	5.792 ±0.97%[b,c]		4.112 ±1.12%[b,c]

[a] Isotope ratios determined in duplicate.

[b] Errors estimated as plus or minus the range of individual aliquots for each concentration level.

[c] Sampling is the dominant error for the results in Exercise 4. The 95% L.E. for the isotope dilution analysis of calcium is ±0.20%, as shown in Exercise 2.

Exercise 5 (Table 10)

The results for this exercise represent the quintessence of our efforts toward an absolute serum analysis, for several reasons.

The Center for Disease Control supplied bovine serum samples that were processed for homogeneity and protected against bacteriological action by the addition of inhibiting agents. Sampling error therefore was not a significant problem in this exercise. An indication of the homogeneity of the samples was observed in the density determinations, which varied only one part in 10^4 over all four lots of serum. In Exercise 3 independent density measurements fluctuated by as much as 0.1% for the individual lots of serum, and by $\sim$0.5% between lots.

In any development work the opportunity to refine various aspects of an analytical technique always results in an eventual improvement in the resultant data. Through small refinements in the mass spectrometric technique such as sample loading and more stringent control of isotopic fractionation, and in the chemical techniques such as better blank control and sampling procedures, the total technique was honed to peak performance. These refinements are reflected in the data for this exercise.

Sample groups 621-625 and 605-609 probably exhibit the ultimate accuracy attainable by this technique. In fact, the agreement is actually better than the present overall ability to consistently reproduce the calcium isotope ratios.

Duplicate isotope ratios on outlying values in groups 660-669 and 641-645 reproduced at $\leq\sim$0.25%, which approaches the value for the random isotope ratio measurement error as calculated in Exercise 2. In fact, the 95% L.E. for the data of this exercise was calculated as 0.21%, which is in excellent agreement with the 95% L.E. of 0.20% for Exercise 2. However, the average of the individual aliquots should be closer to the "true" or absolute value, as was the case in

Exercises 1 and 2 where the "true" value was known.

In addition to the above examples of a refined technique, the three ^{44}Ca spike calibrations agreed to <0.08% and the blank correction was <0.03% for all the samples analyzed.

Table 10. Concentration of Calcium in Serum, meq/liter. (Exercise 5.)

Sample		Isotope Dilution
621		3.569
624		3.573
625		3.569
	Average	3.570
605		4.294
608		4.293
609		4.294
	Average	4.294
660		5.015
662		5.023
669		5.032[a]
	Average	5.023
641		5.728
643		5.742[a]
645		5.729
	Average	5.733
	95% L.E.	±0.21%[b]

[a] Isotope ratios determined in duplicate.

[b] Error expressed as 95% L.E. for a single analysis.

VIII. DISCUSSION

Because the candidate method [4] had been so carefully investigated, it was not thought necessary by the Experts Committee to reinvestigate the factors that formed the basis of that work. Therefore, there was no attempt made to restudy such factors as: interferences, reagent purity, recovery studies, comparison with other methods, effect of acids, etc., all of which were adequately covered by Pybus, et al., in their work. Indeed, the conclusion by those authors was that accuracy was limited solely by the achieveable precision of the method. However, during the course of that work it became apparent that many parameters had to be closely controlled and special attention was paid to these during the course of these studies. Perhaps the greatest unknown factor was what contribution introduction of spectrometers (other companies, other models) other than the type and model used by Pybus, et al., would make to the inaccuracy of the system. The other factors are: analytical chemical techniques (aliquoting, pipetting, weighing, etc.), sample condition, quality and preparation of glassware, reagent quality (including water), significant deviations from the protocol, and recording and reporting of data. All these will be discussed in turn.

A. Instruments

1. Balances

Although no data was collected on the actual instruments used, the question was raised during the June 1971 meeting whether all laboratories could weigh accurately to 0.1 mg. No exceptions were noted. With modern balances now readily available to weigh to this degree of accuracy, no further discussion was felt necessary, and it was assumed that all weighings would in fact be so made. Furthermore, the directions for drying the

$CaCO_3$ SRM are given on the Certificate of Analysis and no exceptions to that procedure were noted.

2. Atomic Absorption Spectrometers

There is no evidence that any one of the different instruments used (see table 3) was superior. The most consistent laboratories throughout (Labs A, B, C), in fact, each used a different type and/or model. However, as shall be noted, the operating environment, stability, and reproducibility of the instruments were extremely important. Operating conditions varied, in many parameters, over wide ranges and no discernible effects on overall accuracy are noted.

a. Burner

This was a constant factor. Every laboratory used the Boling type triple-slot burnerhead, and no difficulties were noted except during the course of Exercises 3 and 4 when particulate residue matter from the samples tended to clog the aspirator. Lab D replaced its burner for Exercise 5 and then noted that its calibration curve became essentially linear, a condition not previously obtained. Drafty ambient conditions affecting burner performance undoubtedly played some role in the troubles encountered by Lab F.

b. Oxidant-Fuel

All laboratories used air-acetylene for the oxidant-fuel mixture. The condition of the flame would seem to be not critical. Lab D used an oxidizing flame, the others either a slightly reducing or stoichiometric condition. It is recommended that either of the latter two be used.

Pressure and flow conditions varied over quite a wide range as shown in Table 4. No comments concerning flame conditions or rate of flow were received

with one exception. Lab A said it was important not to vary or to allow variations in flame conditions during the course of a run and recommends a well-designed and stable gas handling system.

c. Stability and Reproducibility of Readout

By this is not meant solely the stability of the readout device alone (usually digital), but the stability and precision of the entire instrumental system as recorded at the output. Through Exercise 3, variations in readout to ± 1% were tolerated, but after the June 1971 meeting, all participants agreed that reproducibility at the ± 0.5% level was necessary, if the overall accuracy goal of ± 1.0% of the true value was to be made achievable. In Exercise 4 and 5, Lab F dropped out of the network when it was evident that the instrument could not meet this specification. From the start, Lab F had difficulties in this regard (see Comments Section). Lab E aborted Exercise 4 for the same reason, although sampling difficulties probably contributed to its difficulties. Upon reloaction of the instrument, with more stable electrical line conditions, Lab E rejoined the network for Exercise 5.

Lab C noted in Exercise 1 that its noise level was high, but made no further comments on the matter and met the precision requirements throughout all five exercises. Lab K noted an improvement in noise levels after optical surfaces were cleaned.

d. Linearity

No special studies were performed to ascertain the linearity of the calibration curves, although comments on non-linearity were received by Labs D and F (see Comments). The procedure calls for bracketing the unknown solutions with the standard solutions whose calcium values are close to that

of the unknown. Linear extrapolation is used and small departures from linearity are disregarded, although each laboratory ascertained the overall linearity of its calibration curve. Effects of non-linearity in the range for which this method is applicable are small compared to other sources of bias.

e. Source

All laboratories used calcium single-element hollow-cathode lamps. Current conditions ranged from 7-15 mA (Table 3). No comments concerning this parameter were received and none were made during the June 1971 meeting. Evidently, the source and source conditions are not a factor affecting accuracy at the level of this method.

It should be emphasized that all participants in the work were completely familiar with atomic absorption spectrometers, could and would recognize instrumental difficulties, and, in some instances, could make the necessary adjustments without help from service personnel. It was also apparent that the instruments--by and large--were kept in optimum operating condition, although Labs E and F had special difficulties that however reflect in no way on the expertise of those investigators. It is safe to conclude that a properly maintained spectrometer of the 3 types used (it is not possible to speculate on other instruments) is capable of reproducibility, over several hours, of $\pm$ 0.5% and is certainly not the limiting factor in the accuracy of this method.

B. Reagents

Reagents and their quality are specified in the method. In all instances, standard solutions were prepared using

the NBS-$CaCO_3$ (SRM 915). In any referee method the use of certified Standard Reference Materials is advisable as a bias, in the form of unknown impurities, for example, will distort the results in an unknown way. This SRM, supplied and certified by NBS, is of very high purity and although only certified at 99.9+% purity, it is known that the purity, in fact, is at least 99.99%. Furthermore, metallic impurities are known to be small and no significant error is introduced into the system when SRM 915 is used. One word of caution is in order. The SRM is apt to take up a small amount of moisture rather quickly, especially on warm, humid days. Weighing of the material should be accomplished expeditiously. One investigator (Lab A) recommended that 10 times the amount of material specified in the method be used in making up the standard solutions and that then an additional 10-fold dilution be made. If quick weighing is done, this additional precaution is not considered necessary.

Clinical laboratories very often do not have adequate supplies of high-purity water. It is highly recommended that both deionized and deionized distilled water be available. All laboratories used as a minimum deionized water, but two labs (Lab A and C) had deionized distilled water also available. Proper rinsing of all glassware requires copious amounts of pure water and serious error can result if the calcium contained in detergents and tap water is not completely removed during rinsing.

La_2O_3, available from at least the 3 sources used in this work, was known to contain less than 15 micrograms calcium per gram of the oxide. It is not available as an ACS Analytical Reagent (AR) Grade chemical, and all due precautions should be taken to assure the purity of the La_2O_3 used.

All other reagents used were ACS-AR Grade and are readily available from a number of reputable manufacturers.

C. Glassware

All glassware used in the sampling or dilution processes must meet the NBS Class A specifications. This involves some additional expense, but is required if errors from this source are to be held to less than 0.1%. The stratagem of using the identical 10 ml volumetric pipet throughout for sample, blank, diluent, and standard solutions preparation removes the possibility of variable errors in these steps. There is no question but that the time required to run the procedure is markedly increased by this requirement because of the care that must be taken in assuring no cross-contamination between the various solutions. However, until shown (which is now possible) that errors introduced in using several pipets are small, this requirement must be considered an integral part of the method. In one instance, an entire exercise had to be rerun (Lab B - Exercise 5) because the operator deviated from this requirement and used three different pipets. Other variables were also introduced simultaneously and it is not possible to say what contribution to the 1-2% high results of the discarded exercise were due solely to the variation in pipet specification.

Proper cleaning of glassware is of critical importance. The recommended procedure for cleaning all glassware is taken from the candidate method of Pybus, et al., and is the minimum required. Lab C used a more rigorous procedure, which is described in the Comments Section. The final rinsing in deionized distilled water is highly recommended and is standard practice at NBS when ultra-trace level work is performed.

In one instance an entire exercise had to be rerun (Lab H - Exercise 5) when the first results reported were found to be extremely variable. (Deviations from the ID-MS values ranged from - 2.5% to + 6.4%, well outside the precision previously experienced for that laboratory). Upon a close examination of all steps in the procedure, a white residue

was found on several volumetric flasks and it was also found that the rigorous rinsing steps prescribed had not been followed in detail. A rerun of the exercise brought the results shown in section VII.

D. Aliquot and Dilution Steps

In the first exercise, several labs used a 1 ml pipet and a 1:50 dilution--an option allowed in the candidate method as published. In fact, Lab C used an entirely different dilution procedure, which was their standard practice. For Exercise 2 and in the subsequent exercises, these options were removed and the one 10-ml pipet technique required, and all pipetting and dilutions were performed as required with two exceptions. One exception (Lab B - Exercise 5 is described above. The only other exception (Lab C - Exercise 3) came about when it was found that insufficient sample did not allow rinsings to be performed in the manner being practiced in that laboratory at that time. The protocol used for Exercise 3 was not specific enough in spelling out the rinse procedures and this was corrected in the revised protocol used in Exercises 4 and 5.

Three laboratories called into question the pipet rinsing procedures (Lab A, C, and D) and this step was made more specific in the final protocol used in Exercises 4 and 5. Two minor exceptions to the rinse procedures are noted: Lab A used a $HCl-HNO_3$ (3:1) rinse to improve the drainage of the pipet and Lab F in Exercise 1 used a H_2O rinse instead of diluent. Apparently these deviations were unimportant in introducing any significant bias.

In only one instance was temperature allowed to vary in an uncontrolled manner (Lab B - aborted Exercise 5). All laboratories were instructed to allow solutions to reach room temperature before final dilution to the mark. No exceptions or difficulties were noted in this regard.

E. Sample

As mentioned in section IV-A, the scope of this method does not include sampling procedures. However, certain observations may be made based on this work. The pooled sera for Exercises 3 and 4, when finally received by the cooperating laboratories, were in quite a sorry state. Debris and bacterial growth were plainly evident. In fact, Lab D commented, "We would not have accepted this sample for analysis if it had been presented in our routine workload." Yet, overall it will be seen the results were not biased that strongly. It is felt that pipetting errors on these two samples accounted for most of the bias shown in the results. Furthermore, the ID-MS staff had equal difficulties in their aliquoting as shown by decreased precision in their results. Lab A commented, for example, on Exercise 4 that "precision ranged from 0.3 - 1.7% over 6 days on the Exercise 4 samples." Precision on Exercises 1 and 2 were, as shown, much better and more consistent. Pipetting difficulties were also commented on by Lab C for both Exercises 3 and 4. In the opinion of Lab E, the sample was so bad that Exercise 4 was not run.

It may be asked why the Experts Committee allowed these samples to be run. A conscious decision was made to use these two "bad" sera samples in order to determine how far afield the results would scatter. We may conclude that quite murky, turbid solutions may be handled by the method as long as actual precipitates, debris, or growth are absent.

F. Instrument Reading and Valid Measurement Criteria

In Exercises 1, 2, and 3, the concept of a valid measurement was not incorporated in the protocol. During the discussions of all the participants during the June 1971 meeting this became a major topic. Because there is inherent in the atomic absorption spectrometer the possibility of fairly short time instabilities or drifts (electronics, flame, aspirating variations, etc.) during the interval

between readings, it was agreed to set up criteria for accepting or rejecting data. These criteria are set forth in the final method and have the effect of limiting the instrumental errors to ± 0.5%. As many of these instabilities are essentially random in nature, the requirement that the reported result be based on the mean of ten readings reduces this error even further.

Although the laboratories were requested to supply the number of sets of data rejected, only Lab C provided this information and then only for Exercise 4 (see Comments Section). It is suspected that some of the difficulties may have been due to the poor nature of the sample which may have caused aspirator clogging. It would have been extremely interesting to have had this information for Exercise 5.

With the one exception of Lab D in Exercise 4, all laboratories used the ten replicate reading for Exercises 4 and 5.

After Exercise 3 at least one laboratory (Lab C) did not feel that the protocol was clear or definitive enough on defining the run order sequence of standards and samples (for more than one unknown sample as was the case throughout). This was clarified in the final protocol.

G. Deviations from the Pybus, et al., Method

As explained earlier, the candidate method was considered to have been well-thought out enough and studied so that extensive investigation in this study was not necessary. Several options do exist in the candidate method. Some of these have now been removed, others remain because no effect could be attributed one way or the other. Options in the original method now removed include: ten replicate reading requirement; 10:500 dilution factor; and use of one 10 ml pipet. Exceptions made during the course of this work have already been discussed.

Options remaining are principally the use of Sr as an internal reference, and the addition of a Mg standard (very often in clinical use Ca and Mg are determined together).

Laboratories B, H, and K used the Sr internal reference option throughout, the remaining labs did not. There is no evidence from the results that an internal standard improves the accuracy of the method.

Magnesium standard was used by Labs B and C, the others did not. No major effect can be attributed to this factor.

Lab B showed in Exercise 1, as the result of some additional work, that the addition of Na and K was necessary in all solutions--blank, diluent, standards, and unknowns. Exercise 1 samples did not contain Na and K; (it was thought to test the labs for precision primarily on the simplest possible solutions). Lab B commented that if the standards contained Na and K while the test samples did not then results should be about 1% high from the true value. As can be seen for the results from Exercise 1, this conclusion is not borne out. However, it was never ascertained how many of the laboratories added Na and K to the test samples as this point was well known to them. In any case, this question is rather academic because the final method is designed to measure calcium in serum or solutions containing physiological amounts and proportions of Na and K.

Significant deviations from the procedure other than those already noted include:

1. A substantially different calibration procedure (Lab D - Exercise 1). Lab D followed the protocol exactly in Exercise 5.
2. No blank solution used (Lab D - Exercise 1). The argument may be made that an automatic blank correction is made if measurements are not made on an absolute basis, but relative to the standards as both solutions contain the same chemicals and diluent. This may be true, but the value of measuring the blank is to ascertain the magnitude of the calcium blank that should be low relative to the concentrations being sought as blanks are notoriously variable in nature and may result in large biases (often random and unknowable).

3. In the comments and suggestions accompanying the first protocol, the point was strongly made that a sufficient block of time apart from other responsibilities should be made available to the analyst during the running of each exercise. Only Labs A, B, and C were able to or did do this consistently throughout. The other labs are all operating clinical laboratories and this work was scheduled to fit into the routine work load. It is difficult to assess the overall contribution to accuracy of this factor. It will be noted, however, that these three laboratories performed most consistently throughout these studies.
4. One minor point for the record--Lab E in Exercise 1 used a procedure based on another method, but then followed the prescribed protocol thereafter.

H. Discussion of General Points

It would be a mistake to assume that this study was accomplished under ideal conditions and that all parts of the network were under strict control at all times in all exercises. A careful reading of the Comments Section shows how many unplanned variables actually entered the system. Only in Exercise 5 did the network operate in what could be considered adequate control. In the two instances in Exercise 5 where significant deviations were introduced (Lab B and H), quick recognition of these deviations was made by those laboratories and complete reruns were made according to the protocol. Many initial departures were made because there was not a clear understanding in the minds of all the cooperating investigators as to what was allowable and what the underlying philosophy of operation was. It was only after the June 1971 meeting that the point of "no deviation from the protocol" was fully accepted. This study, therefore, should be considered in no small part to be educational in

nature. Scientists in clinical chemistry need to be made aware of and then to accept the importance of properly developed referee methods if a meaningful measurement system is to be realized in practice. It is hoped that the present study will serve as a model for the development of much needed additional referee methods, recognizing the many areas where mistakes in this study were made. Some recommendations to bring about more ideal conditions for future studies are given in section X.

It is readily apparent that the validity of this work rests on the thesis that the isotope-dilution mass-spectrometry reference-line method, against which all the data was referred, is, in fact, a method of proven accuracy. The data and evidence in section VII strongly support this thesis. In fact, the Experts Committee purposely chose calcium as the first referee method to be developed because of the very strong competence of NBS in this area. The question naturally arises--how is a referee method developed when such a base-line reference method is not available? There is no doubt but that the ferreting out of systematic biases then becomes a much more complex and time-consuming task. In such instances, the preliminary studies and research that went into the candidate method must be of the highest caliber--interferences studied, specificity and purity of reagents evaluated, instrumental parameters and errors elucidated, etc. These questions should be fairly well-settled before the method is to be considered a candidate for the development of the referee method. It is evident from this work that only a relatively few variables can be handled within any reasonable expenditure of time, effort, and money. Screening and evaluation of candidate methods by the Experts Committee are thus seen to be of critical importance for future work. When an independent reference-line method is not available, then consensus among experts, but based on

studies done under controlled conditions, is the only viable alternative to the methodology used herein. In this regard, the professional societies, through their standards committees and experts panels, can and should play a leading role.

The question may be raised, "If a reference-line method is already available for use, why not make it the referee method?" In principle there is no reason why this could not be so, but in practice, the reference-line method may have limitations that preclude its adoption as the referee method. This is certainly the case in this instance. There are few, if any, clinical laboratories in the world that have either the equipment or trained personnel to use the ID-MS method described herein; nor would it be practical to suggest its incorporation into the clinical milieu. Unfortunately, it is often true that the most accurately based analytical systems are also those most expensive, both in equipment and personnel.

Although the documented evidence in this study will not rigorously bear out in all instances the following conclusions, there is little doubt of the validity of these assertions:

1. The development of a referee method, given a well-studied candidate method, is essentially an exercise in analytical, not clinical chemistry. The quality, training, background and motivation of the participants, especially those who do the actual laboratory work, must be of the highest caliber. The educational implications for clinical chemistry are obvious. It probably is not fair to ask fully committed and very busy technologists to participate during the course of the work. Motivation will probably be lacking, unless they were involved in the studies from the start and feel they have a personal stakes in this work.

2. Time must be made available for the participating scientist. Referee development work cannot be considered an adjunct or auxiliary responsibility. Again this requirement militates against using cooperating laboratories that are fully committed to high-load daily routines unless no other alternative is available. It should not be concluded that because several of the laboratories in this work are "working" laboratories, that their results were affected adversely because of this pressure. These laboratories produced the results they did in spite of the daily work load. It would not be fair to ask these laboratories to continue on a voluntary basis into the indefinite future. Resources must be made available so personnel dedicated solely to future work can be freed of routine responsibilities--in other words, research in these laboratories should be independently supported.
3. Funding in support of the cooperating laboratories for future work is essential. Only when the work is done in a business like, contractual mode can the necessary controls be instituted and maintained. In this work, the question of controls and direction is academic. All laboratories gave freely of their time and accepted direction in a truly remarkable cooperative manner. There were no prima donnas, no quibbling over details, only valid criticism and valuable suggestions. In the more complex studies to follow, these conditions may not apply, and then it will be important to have the central control that is only possible when the work is funded under a contractual arrangement.

IX. CONCLUSIONS

A referee method for the determination of calcium in serum and/or solutions containing physiological amounts of sodium and potassium has been tested and validated. When the protocol given in section IV is followed in every detail, subject also to the many provisos and conditions listed in the Discussion Section, then this referee method will give the concentration of calcium accurate to within ±2% of the "true" or absolute value in the range from 3 to 6 milliequivalents calcium per liter of sample.

Inherent in the referee method are the conditions and findings of the basic candidate method of Pybus, et al., upon which this work was based. This is especially true for interferences and matrix effects that were not studied in this work. Significant departures from the method with regard to these factors may lead to substantial systematic errors. Both human and bovine pooled serum and aqueous solutions containing physiological amounts of sodium and potassium may be analyzed with assurance by the referee method. The effect on accuracy of scaling down the sample size or using a less than 10:500 dilution has not been tested although some additional work (not reported herein) indicates that smaller sample sizes and lesser dilution factors might be tolerated.

It is certain that the spectrometer used must be in capable, knowledgeable hands, in compatible surroundings, and in tip-top operating condition, especially with regard to stability and reproducibility. Furthermore, no deviation from the measurement reading procedure can be tolerated, as the error from this source alone is likely to approach $\pm$ 0.5%.

Finally, the evidence strongly supports the contentions made in the Discussions Section that the method must be developed by scientists well trained in basic analytical

chemistry techniques and who are, in addition, meticulous throughout and highly motivated.

Note Added

At the conclusion of Exercise 5, it was thought by the Experts Committee to be an interesting experiment to send samples of the same serum utilized in Exercise 5 to several commercial and hospital laboratories.

The results of this small survey are given in Appendix 6. Because of the limited laboratory sample size and the lack of rigorous control, we recommend that no far-reaching conclusions be made based on this work alone. However, the results are illuminating.

It should be noted that one laboratory, Lab R, produced results consistent with the known values. Here, there is the possibility of examining its procedures to see if a faster, less time-consuming referee method can be developed.

X. RECOMMENDATIONS FOR FUTURE REFEREE METHOD DEVELOPMENT STUDIES

To the best knowledge of all concerned in this work, this study is the first made to bring about in a systematic way a referee method in the field of clinical chemistry. It is quite natural that, ex post facto, one now sees the many weaknesses, unnecessary work, inconsistencies, false starts, etc., that might have been avoided through better planning, tighter scheduling, more explicit directions, and closer coordination. In a real sense this study should be considered to have been part of an educational process from which future work will surely benefit. Taking advantage then of hindsight, several recommendations are worth listing.

A. Preplanning Session

No future referee method development study should start until all participants have agreed to meet in person at some agreed on location. Prior to this session, the Experts Committee should have selected, after some considerable study, the candidate method, have set the accuracy goal, and agreed on the factors most likely to be the principal contributors to the systematic errors. The Experts Committee should have prior to this meeting also have met with the Coordinator and statisticians to agree on the experimental design, statistics, etc.

In this study, the meeting of the participants did not take place until after the third exercise. It was then that many details were agreed on that should have been ironed out prior to any of the laboratory work. For example, it was only at that meeting that the "non-deviation from protocol" principle was really accepted.

B. Coordinator

The coordination of the work, gathering the data, sending out the samples, establishing and maintaining communications, and a hundred small details are in the

hands of the project Coordinator. In this work the Coordinator was J. Paul Cali whose primary duty at NBS is the management of the Standard Reference Materials program. Most of the inefficiency, delays in scheduling, etc., can be attributed to the fact that the Coordinator role was secondary to his primary responsibilities, and consequently often the coordinating duties had to take second place. As this work is of the highest importance to the Nation, a full-time Coordinator is highly recommended for future work, especially if the tempo of this work is to be stepped up and increased--both highly desirable goals in view of the vast amount of referee development work to be done.

If additional financial support is made available (next paragraph), then the Coordinator could also act in the role of project manager for: (1) Establishing and monitoring contractual arrangements; (2) Evaluating proposals, and (3) Visiting proposed cooperating laboratories to establish availability of qualified personnel, equipment, etc. These additional tasks would certainly warrant a full-time Coordinator.

C. Financial Support

Without going into great detail, the true cost of developing this referee method was, conservatively, $120,000. Of this amount, probably $40-50,000 of free service was contributed jointly by the Experts Committee and cooperating laboratories. Approximately $40,000 was contributed by NBS (again free to the project) in the salaries and overheads of the Coordinator, his staff, and NBS statistical services. Only $40,000 in direct dollar appropriations was ever charged to this work at NBS, primarily for the ID-MS work.

It is highly recommended that future referee development studies be fully funded, especially in support of

the cooperating laboratories. If this support can be made available, it will accomplish several important ends. It will allow: Fully-qualified and well-trained scientists to be recruited and paid to participate, as opposed to the present situation where scientists must be taken away from other pressing duties; (2) Under contractual agreements, greater control, flexibility, and scope of the work--necessary reruns, more exercises, more comprehensive study of the variables, non-interference by routine work, etc.; (3) Tighter and more intensive scheduling--more work in a shorter duration; and (4) Purchase of necessary supplies, updating or renovating of new equipment.

D. Experts Committee

If financial support can be found, then it is reasonable to expect a more concentrated, extensive study to be made by a paid Experts Committee, rather than one that works completely on a voluntary basis. In this work, it was fortunate that a very well-studied candidate method was already available. In the future this may not be the case, and then the Experts Committee might have to conduct, or have done under their guidance, laboratory work to establish the adequacy of the candidate method.

As has been explained earlier, the availability of an independent reference line method against which the candidate method can be tested is most desirable. An Experts Committee who could spend a fair amount of time on this work, and that implies financial support, could encourage and support developments in this area. For example, studies of isotope dilution techniques using radio-tagged cholesterol or bilirubin would seem to be a subject worthy of support and might provide the independent reference-line method that is not now available for these classes of clinically important substances.

An undertaking of this magnitude involves many people, and, in this instance, many different organizations. It is not possible to give personal credit in every instance because many are not known to the author by name. Valuable advice and expert guidance were received from the Experts Committee--Drs. George N. Bowers, Jr., Donald S. Young, and Nathan Radin. Suggestions concerning some overall philosophy were made by Dr. Russell J. Eilers.

Our deepest thanks are due the cooperating laboratories and the principal investigators whose names and affiliations are given in Section II-A. Without their whole hearted cooperation and enthusiam this work would not have been possible at all. At the same time we acknowledge the contributions of all the scientists and technologists at these laboratories who participated in the actual work.

Special thanks is given to Hartford Hospital (Dr. George N. Bowers, Jr.) and to the Center for Disease Control (Dr. Joseph H. Boutwell, Dr. David D. Bayse and Ms. Dorothy S. Lewis) for the preparation and contribution of the pooled sera used in Exercises 3, 4, and 5.

At NBS many scientists participated in the work: (1) For the isotope-dilution mass-spectrometry reference line studies under the direction of William R. Shields, we thank Ernest L. Garner (mass spectrometry) and Keith M. Sappenfield, Larry A. Machlan, Thomas J. Murphy, and John R. Moody (isotope dilution, preparation of unknown solutions). (2) For reduction of data and assistance with the statistical analysis thanks is due Miss M. Nancy Steel. All the isotope-dilution mass-spectrometry work and the NBS participation as one of the cooperating laboratories was performed in the Analytical Chemistry Division. The Chief of that Division, Dr. W. Wayne Meinke, wholeheartedly supported this effort and always gave this work the highest priority, even in these

times when competition for scarce resources is great. We acknowledge and thank him for this support.

The staff of the Office of Standard Reference Materials helped in many ways during the course of the work: preparation of the data recording sheets, correspondence with the cooperating laboratories, mailing of the samples, etc. Special thanks is due Mrs. Suzanne C. Love for keeping many details straight, and for filing of the incoming data. For the typing of the manuscript thanks are due to Mrs. S. Love and Mrs. V. Davis.

REFERENCES

[1] Eisenhart, C., Realist Evaluation of the Precision and Accuracy of Instrument Calibration Systems, J. Res. Nat. Bur. Stand. (U.S.), 67C, No. 2, 161-187 (Mar.-Apr. 1963)

[2] Cali, J. P., Accuracy in Clinical Analytical Methods and Its Attainment (In publication)

[3] NBS Clinical Laboratory Standards, SRM Brochure (1971) Office of Standard Reference Materials, Washington, D.C. 20234

[4] Pybus, J., Feldman, F. J. and Bowers, G. N. Jr., Measurement of Total Calcium in Serum by Atomic Absorption Spectrophotometry, with Use of a Strontium Internal Reference, Clin. Chem 16, 998-1007 (1970)

[5] Bowers, G. N., Jr., et al, The Preparation and Use of Laboratory Control Materials for Monitoring Precision in the Clinical Laboratory, Standard Methods of Clinical Chemistry, Vol. 7 (in preparation).

[6] Trudeau, D. L., and Freier, E. F., Determination of Calcium in Urine and Serum by Atomic Absorption Spectrophotometry, Clin. Chem. 13, 101 (1967)

[7] Shields, W. R., Editor, Nat. Bur. Stand. (U.S.), Tech. Note (in preparation).

[8] Garner, E. L., Machlan, L. A. and Shields, W. R., Nat. Bur. Stand. (U.S.), Spec. Publ. 260-27, 150 pages, U. S. Government Printing Office, Washington, D.C., 20402 (1971).

[9] Catanzaro, E. J., Champion, C. E., Garner, E. L., Marinenko, G., Sappenfield, K. M. and Shields, W. R., Nat. Bur. Stand. (U.S.), Spec. Publ. 260-17, 70 pages U.S. Government Printing Office, Washington, D.C. 20402 (1970).

[10] Catazaro, E. J., Murphy, T. J., Garner, E. L. and Shields, W. R., J. Res. Nat. Bur. Stand. (U.S.), 70A (Phys. and Chem.) 6, 453-458 (Nov.-Dec. 1966).

[11] Shields, W. R., Murphy, T. J. Garner, E. L. and Dibeler, V. H., J. Am. Chem. Soc., 84, 1519-1522 (1962).

[12] Shields, W. R., Murphy, T. J. Catanzaro, E. J. and Garner, E. L., J. Res. Nat. Bur. Stand. (U.S.), 70A (Phys. and Chem.) 2, 193-197 (Mar.-Apr. 1966)

[13] Shields, W. R., Murphy, T. J. and Garner, E. L., J. Res. Nat. Bur. Stand. (U.S.), 68A (Phys. and Chem.) 6, 589-592 (Nov.-Dec. 1964).

[14] Catanzaro, E. J., Murphy, T. J., Garner, E. L. and Shields, W. R., J. Res. Nat. Bur. Stand. (U.S.), 68A (Phys. and Chem.) 6, 593-599 (Nov.-Dec. 1964).

[15] Catanzaro, E. J., Murphy, T. J., Garner, E. L. and Shields, W. R., J. Res. Nat. Bur. Stand. (U.S.), 73A (Phys. and Chem.) 5, 511-516 (Sept.-Oct. 1969).

[16] Shields, W. R., Garner, E. L. and Dibeler, V. H., Jr. Res. Nat. Bur. Stand. (U.S.), 66A, (Phys. and Chem.) 1, 1-3 (Jan.-Feb. 1962).

[17] Catanzaro, E. J., Murphy, T. J., Shields, W. R. and Garner, E. L., Res. Nat. Bur. Stand. (U.S.), 72A (Phys. and Chem.) 3, 261-267. (May-June 1968).

[18] Shields, W. R., Editor, Nat. Bur. Stand. (U.S.), Tech. Note 277, 99 pages, U.S. Government Printing Office, Washington, D.C. 20402 (1966).

[19] Shields, W. R. Editor, Nat. Bur. Stand. (U.S.), Tech. Note 426, 53 pages, U.S. Government Printing Office, Washington, D.C. 20402 (1967).

APPENDIX 1. NBS STANDARD REFERENCE MATERIALS and REFEREE METHODS DEVELOPMENT

A Referee Method may be defined as a measurement method or technique designed to determine one or more characteristics or properties of a substance to a well-defined and achievable accuracy for a specific or well-defined use. In order to achieve the desired accuracy the method or technique must be capable of study to determine the systematic errors or biases inherent in all measurement systems. For many systems this study is most expeditiously and effectively carried out when a Standard Reference Material (SRM) is available.

An SRM is by definition a well-characterized material whose certified properties are accurately known within some stated limit(s). Through the use of the Standard Reference Material, systematic errors in the method under development may be discovered and eliminated, or if the latter is not possible, then corrected. If this process is carried out in several well-qualified laboratories, by a number of scientists working independently--except for the common use of the identical SRM and identical specimens--then the emergence of a Referee Method becomes possible. When many workers obtain equivalent accuracy with the method, then by consensus among qualified experts we may say that a Referee Method has been achieved.

Throughout this process one sees the central role of the SRM. It is through the mechanism of the SRM that a useful transfer of technology occurs. NBS by its very mission and charter has been charged with this transfer of technology and because it is the primary source of SRM's in the nation, the Bureau can play a significant role in the aforementioned process. NBS, through its Institute for Materials Research,

and its Office of Standard Reference Materials is now ready to implement this process where needs are urgent, identifiable, and in the national interest.

Several ingredients are necessary if this process is to be successful:

1. The undertaking must have the interest and active backing of national, and in some cases international, professional, industrial, or standardizing societies, groups or bodies. The participation of these bodies in defining the problem to be studied, deciding on the experimental parameters, guiding the cooperating laboratories, and evaluating the results should be delegated to a small (3-7) group of well-qualified experts, drawn from the interested societies and groups.

2. The achievement of accuracy in the determination of the characteristics or properties of important classes of substances must be the overriding goal, and is best reached through the development of a referee method based based on an SRM.

3. All participating elements in the referee process must be in quality control at all times. This implies tight statistical control and well-designed and executed experimental techniques.

4. The dedicated commitment of a number of well-qualified laboratories in support of the "round-robin" development and testing procedures is paramount.

5. When the accuracy goals of the method have been achieved, NBS will publish the method as a "Referee

Method" in the publication series reserved for Standard Reference Materials (NBS Special Publication Series 260). Simultaneous publication in professional society journals, other government publications, or standards writing organs is not precluded, and indeed, will be encouraged. These methods as developed will also be referenced in the SRM catalog alongside the pertinent SRM.

NBS participation in this process, when all the necessary conditions listed above have been assured, will consist of all or some of the following (as appropriate):

1. Provide the appropriate SRM to participating laboratories.

2. Provide the statistical design and experimental controls for the various stages of the experimental phases of the work.

3. Provide, where technically feasible, samples of known composition whose properties have been determined at NBS by an independent method of previously ascertained accuracy. (NOTE: In many, indeed in most, cases an accurate independent method is probably not available or if so, too complicated, complex, or expensive to be considered as the candidate referee method, itself).

4. Collect, collate, and statistically evaluate the cooperating laboratory results and report these to the experts team.

5. Participate as one of the cooperating laboratories.

6. Coordinate the activities of the cooperating laboratories, internal NBS work, and the experts team.

7. Publish the results and reference methods in the NBS 260 Special Publication Series.

J. Paul Cali, Chief
Office of Standard Reference Materials
National Bureau of Standards
Washington, D.C. 20234

APPENDIX 2

U. S. Department of Commerce
Maurice H. Stans
Secretary

National Bureau of Standards
A. V. Astin, Director

Certificate of Analysis

Standard Reference Material 915

Calcium Carbonate

This Standard Reference Material is certified for use in the calibration and standardization of procedures employed for the determination of calcium in clinical analysis. The sample consists of highly purified calcium carbonate, and chemical assay as well as analysis for specific impurities indicate that the material may be considered to be essentially pure.

Purity	99.9+ percent
Water	0.01 ± 0.005 percent

Replicate samples taken from a randomly selected region of the undried material were assayed by a coulometric acidimetric procedure. The results from nine independent determinations, based on expression of the assay as calcium carbonate, indicate a purity of 99.99+ percent with a standard deviation of 0.003 percent. Samples equilibrated at a relative humidity of 90 percent and assayed by this coulometric procedure showed a maximum moisture adsorption of 0.02 percent as compared to samples that were dried for 6 hours at 210 °C. The moisture content, similarly measured, on samples equilibrated at 75 percent relative humidity, was found to be 0.01 percent. The water content was determined by the Karl Fischer method.

The calcium carbonate used for this Standard Reference Material was obtained from the J. T. Baker Chemical Company, of Phillipsburg, New Jersey. Analyses were performed by C. E. Champion, E. R. Deardorff, G. Marinenko, O. Menis, T. C. Rains, T. A. Rush, W. P. Schmidt, B. F. Scribner, V. C. Stewart, J. K. Taylor, and D. W. Vomhof.

The overall direction and coordination of technical measurements leading to the certification were under the chairmanship of R. Schaffer.

The technical and support aspects concerning the preparation, certification, and issuance of this Standard Reference Material were coordinated through the Office of Standard Reference Materials by J. L. Hague.

Washington, D. C. 20234
March 4, 1969

W. Wayne Meinke, Chief
Office of Standard Reference Materials

APPENDIX 3 - NAMES AND ADDRESSED OF PRINICIPAL INVESTIGATORS AND COOPERATING LABORATORIES

1. Dr. George N. Bowers, Jr.
   Co-Director
   Clinical Laboratory
   Hartford Hospital
   Hartford, Connecticut 06115

2. Dr. S. Raymond Gambino
   Professor of Pathology and Director
   of Clinical Chemistry Laboratories
   College of Physicians and Surgeons
   of Columbia University
   Department of Pathology
   630 West 168th Street
   New York, New York 10032

3. Dr. Alan Mather
   Chief
   Clinical Chemistry Section
   National Center for Disease Control
   1600 Clifton Road, NE
   Atlanta, Georgia 30333

4. Dr. John C. J. Porter
   Biochemists and Chief
   Department of Clinical Biochemistry
   Toronto General Hospital
   Toronto 2, Ontario Canada

5. Mr. Theodore C. Rains
   Research Chemist
   Analytical Chemistry Division
   National Bureau of Standards
   Chemistry Building, Room B226
   Washington, D. C. 20234

6. Dr. Royden N. Rand
   Assistant Chief
   Chemistry Division
   Hospital of the University
   of Pennsylvania
   William Pepper Laboratory
   3400 Spruce Street
   Philadelphia, Pennsylvania 19104

7. Dr. Leonard Sideman
Chief, Biochemistry Section
Department of Health
Division of Laboratories
2100 West Girard Avenue
Philadelphia, Pennsylvania 19130

8. Dr. Donald S. Young
Chief, Clinical Chemistry
Department of Health, Education,
and Welfare
Public Health Service
National Institute of Health
Building 10, Room 4N 309
Bethesda, Maryland 20014

APPENDIX 4. Preparation of Serum Calcium Control for the National Bureau of Standards Calcium Standardization Program

by

David Bayse, Ph.D.
Dorothy Sue Lewis, B.S.

This calcium control was formulated as follows:

1. Three liters of bovine serum were dialyzed against saline for 24 hours. The saline was changed twice during this period.

2. This dialyzed serum was assayed by calcium atomic absorption and was found to contain 1.39 meq/l calcium.

3. This dialyzed bovine serum was mixed with undialyzed bovine serum with a calcium analysis of 4.69 meq/l. Eight liters of such a mixture were prepared with a final calcium concentration of 3.6 meq/l. This pool constituted both the low pool and the basic matrix for concentration of this study. Antibiotics were added to this matrix:

   672 mg of streptomycin, USP
   1,600,000 units of penicillin, USP
   20 mg of fungizone, USP

4. To the 2.93 liters of the low calcium pool were added 23.2 mls of calcium stock solution (402.2 meq/l calcium). This solution was allowed to equilibrate for 24 hours and was stack filtered using Millipore filters: prefiltered - 1.5 microns, 0.8 microns, 0.6 microns, 0.45 microns, 0.3 microns, 0.22 microns. The final concentration of this solution was 5.69 meq/l.

5. The two intermediate pools were prepared from mixtures of the high and low calcium pools using a Mettler P10 balance. One of the two intermediate pools consisted of 1/3rd of the high pool mixed with 2/3rds of the low pool by weight, with a resulting calcium analysis of 4.29 meq/l. The second intermediate pool consisted of 2/3rds of the high pool and 1/3rd of the low pool by weight, and the resulting calcium was 4.95 meq/l.

6. The material was then sterile filtered using a 0.22 micron Millipore filter and 50 mls were dispensed into 60 ml sterile Wheaton bottles. This material was tested for sterility by microbiological examination.

APPENDIX 5
Sheet 1

Analyst_____ Date Run___ Analysis_____

Exercise No. __ Run No. __ Laboratory Code __ Instrumental Run Time: Start ____ AM PM Completed____ AM PM

Sample Number (Run in order given)	Dilution				Dilution Factor Z*	Absorbance	Concentration from Calibration Curve (µg/ml)= C	Final Concentration (C·Z)/20.04 meq/l	FOR NBS USE
	A_1	V_1	A_2	V_2					

For Comments or Additional Data, use reverse side.

*Dilution Factor = $\frac{V_1}{A_1}\frac{V_2}{A_2}$: where A_1 = aliquot from NBS sample, in ml; A_2 = aliquot from V_1, in ml; V_1 = volume to which A_1 was diluted, in ml; V_2 = volume to which A_2 was diluted, in ml.

APPENDIX 5
Sheet 2

Analysis_____

Exercise No._____ Run No._____ Lab Code_____

NBS Sample #_____	NBS Sample #_____	NBS Sample #_____	NBS Sample #_____
Absorbance Values	Absorbance Values	Absorbance Values	Absorbance Values
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

NBS Sample #_____	NBS Sample #_____	NBS Sample #_____	NBS Sample #_____
Absorbance Values	Absorbance Values	Absorbance Values	Absorbance Values
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

APPENDIX 5
Sheet 3

Exercise No._____ **Run No.**_____ **Method**_____

Date received_____ **Date reported**_____

Instrument (Company)_____ (Model #)_____ (Serial #)_____

Wavelength,_____ Å **Slit width**_____ µm

Source, (single or multielement) hollow-cathode lamp

Current,_____ **mA**

Burner,_____ **(type)**

Oxidant,_____ **flow rate**_____ **l/min.**

Fuel,_____ **flow rate**_____ **l/min.**

If flow rate unknown give: Oxidant_____ **(PSI);**

fuel:_____ **(PSI)**

Flame condition, (oxidizing, reducing or stoichiometric)

Time constant,_____ **sec. Scale expansion**_____ **X.**

Readout obtained with (recorder, digital or_____**)**

Standard stock solution,___ **mg/l, prepared from** _____.

Calibration curve, working range_____ **to**_____ **µg/ml.**

Laboratory Ambient temp_____ °C to _____ °C (over course of entire exercise)

Interference Filter (If used): _____ type; _____ nom. wavelength

Other comments: (Note especially deviations from Bowers method)

APPENDIX 6 - RESULTS FOR CALCIUM (USING EXERCISE 5 SERUM) FROM 9 LABORATORIES USING ROUTINE METHODS.

(Donald S. Young)

The performance goals of this study were achieved with the Reference Laboratories only after five rounds of analysis of samples. These laboratories were selected in the first place for their reputation for high quality work and the concern of their directors for accuracy and precision in analysis. All of these laboratories used atomic absorption spectrophotometry for measurement of calcium. We felt that it would be interesting to observe the performance of other laboratories without the intensive preparation that was necessary to achieve the satisfactory performance eventually accomplished by the Reference Laboratories. We initially made the assumption that atomic absorption was the best method for assaying calcium in physiological fluids, without evaluation of other techniques. Therefore we elected to send the same samples that were analyzed in the fifth round by the Referee Laboratories to several hospital and commercial laboratories for analysis by their routine methods. The type of laboratory and technique used is summarized in Sheet 1.

Each of the four serum samples analyzed by the Reference Laboratories in the fifth round of the study was divided into two parts (labeled 1 and 2). These were sent as samples of unknown concentration to each of the participating laboratories. The staffs of these laboratories were requested to analyze the samples without special handling. None of the analysts was aware that the two samples labeled with the same code letter were aliquots of the same serum. The aliquots were prepared in the Clinical Pathology Department, National Institutes of Health. The aliquots were made from several bottles of the serum used in the fifth round of the survey of the

Reference laboratories. The aliquots were dispensed into acid-washed Vacutainers and stored at 4°C for less than one week before analysis. The aliquots were clear and appeared to be homogeneous. However, the Director of laboratory P initially questioned the homogeneity of the samples. The Director of laboratory X noted the presence of a slight deposit in the samples that he received. None of the other laboratory Directors commented on the samples.

The results of the analysis of the samples by the different laboratories are presented in Sheet 2. Within all laboratories for each technique there is good consistency between the duplicate measurements. The largest discrepancy observed was 0.2 meq/l. One laboratory (laboratory R) found no difference between the paired results. However in those laboratories in which two methods were used there was poor agreement between the mean results by the two techniques. In laboratory P, although there was internal consistency when the samples were analyzed at night and during the day, the mean values differed. The laboratory Director believes that the analyst at night failed to mix the samples adequately prior to analysis.

For each sample there was a considerable spread of results when the performance of all laboratories was compared. Thus results for sample A varied from 3.8 to 4.7 meq/l, for sample B from 3.25 to 4.3 meq/l, for sample C from 4.85 to 6.0 meq/l, and for sample D from 4.25 to 5.6 meq/l. The spread of results in each case brackets or exceeds what would normally be considered as the normal range of results of a healthy population. With the diagnosis of a parathyroid tumor still determined to a large extent by the concentration of calcium in a patient's serum several unnecessary and dangerous operations would have been indicated, on the basis

of the results obtained by the different laboratories.

There was no consistency between the laboratories using the same technique suggesting that factors other than methodology are responsible for the variation in results. That it is possible for excellent results to be achieved under routine conditions is indicated by the performance of laboratory R. All the duplicate values were identical and all agreed with the isotope dilution values obtained at NBS and the mean value of the Referee laboratories. While it can not be guaranteed that no special precautions were taken for the analysis of these samples in laboratory R it is gratifying to note its performance because it is one of the biggest commercial laboratories in the country. The impact of its performance is probably greater than that of any other laboratory. On the other hand the performance of another large commercial laboratory (U) was the worst of all the participating laboratories. As a consequence of this study this laboratory has now identified the cause of its poor performance and corrected its deficiencies.

This study has confirmed the widely-held belief that calcium measurements are generally inaccurate and lack consistency from laboratory-to-laboratory. However, the performance of some of the laboratories indicates that good results can be achieved under normal operating conditions. This study confirms our contention that in order to achieve meaningful values in clinical laboratories a national program on a scale much broader than that envisaged hitherto, must be undertaken to educate analysts and examine the cause of the deficiencies in their laboratories.

APPENDIX 6

Sheet 1

TYPE OF LABORATORY AND METHODOLOGY USED

LABORATORY	TYPE	TECHNIQUE
N	Hospital	Atomic Absorption
P	Hospital	EDTA Titration
Q	Hospital	EDTA Titration
R	Commercial	Atomic Absorption
S	Commercial	a) Colorimetric EDTA (SMA 12/60) b) Flame emission
T	Commercial	Fluorometric EDTA (AutoAnalyzer)
U	Commercial	Atomic Absorption
W	Commercial	EDTA Titration
X	Hospital	a) EDTA Titration b) Atomic Absorption

APPENDIX 6

Sheet 2

CALCIUM SURVEY (RESULTS MEQ/L)

(Serum from Exercise 5)

REFEREE METHOD Values	SAMPLE	LAB	LABORATORY P		LAB	LAB	LAB S		LAB	LAB	LAB	LAB X	
		N	DAY	NIGHT	Q	R	SMA 12/60	FLAME EMISSION	T	U	W	EDTA	AAS
4.29	A 1	4.2	4.4	4.65	4.05	4.3	4.43	4.6	4.3	4.45	3.83	4.0	4.3
	A 2	4.2	4.45	4.45	4.0	4.3	4.45	4.7	4.35	4.45	3.85	4.26	4.5
3.57	B 1	3.5	3.6	4.0	3.4	3.6	3.8	4.3	3.65	3.45	3.25	3.41	3.7
	B 2	3.5	3.6	3.8	3.4	3.6	3.8	4.2	3.7	3.25	3.25	3.52	3.85
5.73	C 1	5.5	5.6	5.88	5.85	5.9	6.0	6.0	5.9	4.85	5.33	5.4	5.8
	C 2	5.65	5.55	5.75	5.85	5.7	5.9	5.8	5.9	5.05	5.4	5.46	-
5.02	D 1	4.85	5.0	5.23	4.6	5.0	5.3	5.3	5.2	4.25	4.6	4.98	5.45
	D 2	4.8	4.9	5.33	4.65	5.0	5.2	5.23	5.15	4.25	4.6	4.92	5.6