Pure & Appl. Chem., Vol. 45, pp. 63–68. Pergamon Press, 1976. Printed in Great Britain.

RATIONALE FOR REFERENCE METHODS IN CLINICAL CHEMISTRY

J. PAUL CALI
Office of Standard Reference Materials, National Bureau of Standards, Washington, D.C. 20234, U.S.A.

Abstract—In the field of clinical chemistry evidence is mounting that a reliable and medically relevant structure can be built on the concept of compatibility through accuracy in measurement. By compatibility is meant the ability of all laboratories in a network to achieve on a given sample similar and reliable numerical values for a property under test. It is shown that when all laboratories in a network are making accurate measurements, i.e. free of systematic error and precise, then compatibility automatically ensues. The need for accurate measurement based on medical arguments is given.

In order to build an accurate measurement network, three measurement methodology levels and three types of reference materials are required. A *definitive method* is directly able to realize or to have access to the base or derived units of the measurement system. It depends in part on the availability of *pure reference materials*. (E.g. those supplied by the National Bureau of Standards and called SRM's.) At the next level are *reference methods* and *matrix reference materials*. These methods and materials are more adapted to implementation and use by clinical reference laboratories and the manufacturers of *secondary reference materials*. In turn these materials are used at the local level to control the quality and accuracy of the *routine methods*. How this hierarchy of methods and reference materials is structured is discussed.

The role of the definitive method is discussed in some detail, because of its crucial place in the structure. Potentially useful definitive methods, both for inorganic and organic constituents are outlined. A discussion of reference method developments, per se, is left to the other three authors of this Symposium (qv). Clinical SRM's now available are listed, as well as those now in preparation at NBS. The need for matrix reference materials is stated.

Finally, how this measurement network should be structured and implemented together with suggested organizations responsible for the various levels is discussed.

1. INTRODUCTION

Three recent events in the field of clinical chemistry lead one to believe that a world consensus is rapidly being arrived at concerning the best way to improve and maintain the reliability of clinical measurements. In April 1974, at a meeting of the First European Congress of Clinical Chemistry in Munich, Germany, 18 experts from 10 countries discussed various activities all dedicated to the improvement of measurement in clinical chemistry, stressing in particular the need for reference materials and reference methodology. Within the International Federation of Clinical Chemists an Office on Reference Materials and Methods has been proposed to coordinate international efforts in these areas. Finally, in April 1974, the 27th World Health Assembly directed the World Health Organization to establish an Office on the Standardization of Diagnostic Materials. Such an office is now being made operational in Geneva. It is fair to say that these events are symptomatic of a quiet revolution now occurring on a worldwide basis and all dedicated to the improvement of measurement in clinical chemistry. This revolution is predicated to some extent upon the rediscovery of some quite basic principles now in need of systematic implementation. These principles are: that a major part of clinical chemistry is or ought to be thought of as clinical *analytical* chemistry; that the recognized medical need for reliable quantitation of results is growing rapidly; that large analytical measurement errors can no longer be tolerated either from a scientific or economic viewpoint; that compatible (to be defined) measurement networks are not spontaneously generic, but must be consciously built and nourished. It is to this latter point that this paper is primarily addressed.

2. MEASUREMENT AND MEASUREMENT COMPATIBILITY

2.1 *Measurement*

Measurement in sicence is that process whereby a numerical value is associated with a distinct, specific, and unique property of a material. The magnitude of the number is related to the degree or amount of that property in a particular material or similar class of materials. There are two essential ingredients in this process: a scale, to which the magnitude of the number realized can be related; and, a method whereby the scale is applied in the process to obtain the numerical value in a reliably reproducible manner. Because we are considering here primarily the measurement of the compositions of materials (e.g. cholesterol in serum) our scale will be a material of known composition (e.g. pure cholesterol or a known amount of cholesterol in serum) and our method will be the set of rules or procedures whereby that scale is applied. As will be shown there is not one unique material or one unique method, but rather a hierarchy of these depending upon where in the measurement network these are being applied. In 1939, Shewart[1] described the above two aspects in terms he called the quantitative and the qualitative. The former aspect concerns numbers associated with a pointer reading, a meter, a counter, etc. and in this paper will be associated with a material whose property(ies) have been well-characterized. One class of such materials, now called generically "Reference Materials," are the Standard Reference Materials (SRM's) measured, certified, and issued by the U.S. National Bureau of Standards (NBS). The qualitative aspects Shewart included in what is often called the procedure of method, and in clinical chemistry, the protocol. Included

in this factor are all those things that are used or can affect the course of the measurement process. Among others these are the experimenter or measurer himself, the sequence of operations, the ambient conditions, the apparatus, indeed, every significant parameter that affects the end result.

There is no reason in principle why many different measurement scales could not be used, as long as their relationships one to the other are known. Indeed, historically this has been the case, but the scientific, economic and social benefits from the use of one coherent set of scales is so obvious (at least in the scientific community), that agreement in principle has been reached that all should use the universal system called the Systéme International D'Unités (SI).[2,3]

Having defined and agreed on the units and their derivatives, access to them is provided through highly refined measurement processes—the science of metrology. The transfer of this process throughout a measurement network is often accomplished through the use of artifacts, such as a set of weights, volumetric glassware SRM's, and and like whose magnitudes have been carefully established.

2.2 *Measurement compatibility*

The meaning of measurement compatibility flows directly from the most common usage of the word compatible, namely, the ability to get along well together. If two or more laboratories working independently agree to measure a specific property of a particular lot of stable material, and if each obtains numerical values that agree one with the other, then that network is said to be a compatible measurement network. In real life there will be, even in such a network, slight divergences in the numbers and the question "Agree within what limits?" must be added. In practice these uncertainty limits should reflect the end-use requirements for which the measurement is to be made, and one is then concerned with verifying that the results obtained are compatible within the agreed-on limits.

It is not often realized that measurement compatible networks can be based on many different modes. Three of the most widely used are: (a) the calibration of measuring in struments sent to a well-qualified laboratory and returned to the user;[4] (b) the publication of critically evaluated Standard Reference Data, which, if given with detailed preparation and measurement procedures allows the user to use the data directly or to reproduce the original measurements;[5] and, the provision, directly from a central source to the user, of signals for the measurement of time or frequency.[6] Here, we will examine still another mode, most appropriate to clinical chemical measurements, namely, the achievement and transfer of measurement compatibility by means of reference materials and reference methods of demonstrated and known accuracy.

3. COMPATIBILITY THROUGH ACCURATE MEASUREMENT

3.1 "*True value*" *concept*

Every property of a specific, homogeneous, stable material has a number on some scale that is its actual value. This value is often called the "true value." In a paper by Dorsey and Eisenhart,[7] and highly recommended for all who are interested in the philosophical bases for absolute measurement experiments, the "true value" is called the quaesitum. Intuitively, most scientists would agree that if all laboratories in a measurement network are proceeding in such a way that each operational step in the process was directly traceable to the "true value" (for a given property) and that each step was made without error, other than the measured but irreducible random errors always present, then measurement compatibility in that network exists. This logic is based on the assumption that there can exist only *one* "true value" for the property of the material under examination. Measurements that can be rigorously related to the "true value" will be called *accurate measurements*, and by definition, or through acceptance of the above philosophy or logic, are compatible. Such measurements must agree.

An accurate measurement system must then be free of systematic errors, those errors that would lead one away from the "true value". In addition, however, the measurements must also be considerably more precise (reproducible) than the small and usually unknown systematic errors remaining after all is done. Man being imperfect can never with certainty know he has removed every systematic error; he can only be careful and cautious. Cali and Reed[8] have shown the difficulty of ferretting out systematic errors in an imprecise system.

3.2 *Systematic error*

When systematic errors are present in a measurement process, then the numerical results differs from the "true value" usually, but not always, by a bias of fairly constant magnitude and direction. Biases due to slowly changing phenomena (e.g. humidity) and not recognized as affecting the measurement are especially insidious and may, in fact, often be included in the imprecision statement of the experiments. In an operational sense, if the investigator has available a reference material and a reference method, he can first test his measurement process to see if he can obtain the "true value" into the reference material. Any deviation from that value can be attributed to systematic errors in his (uncorrected) measurement system. He is now in the position of being able to correct these systematic errors because he knows both the direction and magnitude of his systematic errors. When he has properly adjusted his apparatus, procedures, etc. then his subsequent measurements can be said to provide "true values". Furthermore, the value obtained through this process should now be essentially the same as any other acceptable (i.e. accurate) process used to measure the same property of the same material. This treatment is extremely simplistic and the reader is directed to the National Bureau of Standards series on Precision Measurement and Calibration[9] and A Code of Practice for the Detailed Statement of Accuracy[10] for a more complete discussion.

3.3 *Precision*

Precision, or more correctly imprecision, is a measure of the random errors residing in any measurement process. As was commented on earlier, systematic errors much smaller than the random errors cannot easily be detected and then eliminated. Thus, the preliminary attainment of a certain degree of precision is always a prerequisite for an accurate measurement system. *But*, highly precise systems may be often highly inaccurate. This is a real danger and one to be considered at all times.

In practice, the interplay of varying degrees of inaccuracy with varying degrees of imprecision leads to some highly interesting situations. These are discussed in depth by Eisenhart.[11]

3.4 *Other desirable characteristics of an accurate measurement system*

An accurate measurement system must have the attributes outlined above, i.e. freedom from systematic error and precision. In addition, there are other desirable characteristics that from a pragmatic point of view are of considerable importance, and indeed, may be required. These include: specificity, (whose lack constitutes an especially insidious type of systematic error), sensitivity of detection, large dynamic range, ease of operation, high speed, low cost, and several others. Fortunately, the use of reference materials and reference methods to achieve accurate measurement often helps bring about these other characteristics.

4. MEDICAL NEED FOR ACCURATE MEASUREMENT

Cali has discussed previously problems of standardization in clinical chemistry[12] and a systematic approach to accuracy in clinical chemistry.[13] These discussions were based primarily on good measurement considerations and did not touch on the medical needs for accuracy in measurement. Recently, however, the respected pathologist, Roy N. Barnett, M.D., wrote a paper entitled "Accuracy Needs of the Physician for Diagnosis and Treatment".[14] There are four situations cited by Barnett where accuracy is a medical requirement: (a) "to conform with published values when there are accepted levels for separating normal from diseased individuals;" (b) "when there exists a physiologic reciprocal relationship between two or more analytes in the same sample;" (c) "when dosage of medication is predicated on the determined level of some blood constituent;" and, (d) "when metabolic exchange studies serve as a guide to diagnosis and treatment." Certainly, these stated medical needs lend great weight to the rationale for accuracte measurement in clinical chemistry.

5. THE HIERARCHY OF MEASUREMENT METHODS AND REFERENCE MATERIALS

If the thesis is accepted that accuracy in measurement is needed throughout the entire clinical chemistry laboratory network, then it remains to elucidate possible mechanisms to assure its implementation in practice. First, it should be recognized that there exists in all complex measurement networks not one, but a variety of laboratories all in an interacting mode. Second, there exists (or should exist) a variety of reference materials each suited to the particular level of work being carried on. Third, there should be measurement methods available that are suitable for the particular level to which it is to be applied. These relationships are shown in Table 1. The base upon which the entire structure rests, is, of course, the base and derived units of the S.I. Access to or realization of these units is the task of what is now called the *Definitive Method* based in part on the pure reference material (called at NBS the SRM). The next level involves a *Reference Method* and an SRM, but now provided in the actual matrix under measurement. The role of the manufacturer is to use the Reference Method and matrix SRM to assure the accuracy of his product, the quality control sera, reagents, kits, etc. used at the local laboratory level to assure, in turn, the accuracy of the routine of field methods. We have also suggested in the first column various agencies who should take the responsibility for the development of methods and materials appropriate at the level shown. In the last column we give a very rough indication of the degree of accuracy that will be needed at the various levels. If, for example, the medical need at the local level (where it should be generated) is such that the analysis should provide results that are within ±10% of the true value, then each lower level method will require an increased degree of accuracy, in about the amount shown. This comes about because there is always (in practice) a loss of accuracy that occurs during the transfer process throughout the network. It is emphasized that these are approximate and may vary quite markedly with the specific method under development. Each case, with its own pecularities, difficulties, and state-of-the-art will determine the actual accuracy goals to be set.

Table 1. The hierarchy of measurement methods and reference materials

Suggested responsible agency	Method of measurement	Required reference material	Accuracy required (example)
Research and reference laboratories; Instrument and diagnostic material manufacturers	Routine or field	Q.C. sera, kits, reference sera, reagents, etc.	+3–5%
Professional societies, government labs, standard labs	Reference	SRM matrixed	±1–2%
National standards labs, international lab networks	Definitive	SRM pure	+0.2–0.5%

The foundation stone, the base and derived units of the SI

The questions that most often arise when such a scheme as this is proposed is "why is such an involved structure necessary? Why shouldn't the reference method, for example, be also the routine method?" Or, "Why can't the national laboratory supply directly a matrix SRM, or supply Q.C. sera, etc.?" The answer to these and similar questions are based largely on pragmatic and economic considerations. For example, as will be discussed in more detail later on, a definite method usually involves complex, costly, and highly sophisticated instrumentation manned by highly skilled and trained scientific specialists. The instruments and personnel for the development of definitive methods exist, literally, in only a small number of laboratories throughout the world. Or, take reference method development. These methods will in most instances be based on taking a large sample, will require a closely coordinated network of 6–10 laboratories, must have direct access to the definitive method laboratory, etc. The reference method when developed, will usually be time consuming and not adaptable directly to routine use. In short, this enterprise is one where the principle of division of labor makes great sense.

In the ensuing three sections, 5.1–5.3, we will discuss in turn definitive, reference, and routine methods. Emphasis is placed on the first, because reference and routine methodology has been widely covered in the literature. In addition, the subject of reference materials has been widely discussed previously and will not be covered in depth here.

5.1 *Definitive methods and pure reference materials*

Definitive methods have also been called absolute methods. They are absolute in the sense that all significant parameters affecting the final result are traceable by direct experimental evidence to the base and derived SI units. Thus, if mass, time, and temperature are critical parameters in the measurement process then in the exposition of the written method, the scientist will say how each of these has been controlled, how traceability to the base or derived units has been accomplished, how stability with respect to these has been assured, etc. Further, he will be able to give with a high degree of confidence bounds to the limits of uncertainty (i.e. the limits to the systematic errors).

The paper by Churney and Armstrong[15] outlining the calorimetric method used for the measurement and certification of the Benzoic Acid SRM 39*i* is an example par excellence of the definitive method.

The article by Moore and Machlan[16] outlining a definitive method, the determination of calcium in serum by isotope-dilution mass spectrometry (ID–MS), is an example in the clinical chemical field.

At NBS, ID–MS has been used extensively as a definitive method for the accurate measurement of inorganic constituents at the trace levels. The equation which relates concentration to the analytical parameters is:

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

where

W_{sp} = Weight of spike solution, g
C = Concentration of spike, μmol/g 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 of A/B in the spiked sample
M = Atomic weight of analyte
W_s = Weight of sample, g

The power of the ID–MS method rests on 2 aspects. First, all chemical manipulations are done on a weight basis and involve straightforward stochiometric separations, precipitations, etc. to determine W_{sp}, C and W_s. Second, the mass spectrometric determinations involve only ratios and not the absolute determinations of the isotopes involved. Therefore, no instrumental corrections or errors are involved. This, of course, is an oversimplification of the experimental situation and readers are urged to examine the references given above.

The accuracy and reliability of the definitive method is illustrated from recent work at NBS where a reference method for lead in blood is now under development.[17] Solutions of lead at three different concentrations (nominally 4, 2 and 0·7 μmol/l) were prepared by carefully weighing a high purity lead and dissolving in acid. These solutions were given as unknowns to the scientists performing the ID–MS work. The *precision* of the method is shown by the data in Table 2.

High precision, as demonstrated here, is almost always a concommitant of accurate methods.

The *accuracy* of the method can now be demonstrated by comparing the ID–MS results with the weighed in values. The errors associated with the preparation of the sample solutions have been assessed very carefully over a long period of time and are those associated primarily with weighing errors that have been determined to be of the order of a few parts in 100,000 or less. The purity and isotopic composition of the primary "weighed-in" lead is also known to one part in 200,000. The accuracy of results of the ID–MS method for lead are shown in Table 3.

Table 2. Reproducibility of lead determination by ID–MS

Sample No.	Concentration (μmol/l at 23°C)
1	0·68166
2	0·68127
3	0.68190
4	0.68103
5	0.68142
6	0.68137
	0.68137

Mean = 0·68144; Std. Dev. = 0·00030; C.V. −0·044%.

Accuracy when calculated as shown in the table is thus a few parts in 10,000. However, when allowances are made for unknown sources of systematic error, an overall uncertainty of about 1 part in 1000 is claimed for the ID–MS definitive method for lead. Similar results, in terms of precision and accuracy were also experienced when lead in procine blood was determined as part of the lead in blood reference method development.

Obviously, limitations of time, money, technical skills and resources preclude the widespread use of the definitive method. Further, most analytical methods cannot ever be classified as definitive methods, usually because there is no straightforward theory that relates all the experimental variables to the final result. For example it would be difficult, if not impossible, to conceive of a definitive method based on emission spectroscopy simply because the theory that related the energy (or light intensity) in a given spectral line to the concentration of the excited species is much too complex for direct laboratory validation.

5.1.1 *Potential definitive methods for clinical chemistry.* To date only ID–MS has been applied as a definitive method for use in clinical chemical standardization. This research has been well-documented and its history will not be repeated here.[16,18] Currently, this technique is being used at NBS for the other electrolytes (excepting sodium) in serum.

Other techniques of promise and potential as definitive methods for inorganic constituents of interest to clinical chemists are: (a) polarography; (b) activation analysis using chemical separations; (c) and, atomic absorption spectrometry. However, much work must still be done before the "definitive" appellation can be applied with a high degree of certitude.

Table 3. Accuracy of lead determination by ID–MS

Series No.	Lead concentration (μmol/l) Known	Found	Accuracy (found/known)
A	0·68104	0·68142	1·00056
B	2·4080	2·4087	1·00029
C	4.3899	4·3892	0·99984

In the analysis of organic constituents the situation is far less advanced and major research problems must be overcome before any definitive methodology emerges. Techniques holding promise are: (a) organic mass spectrometry based on isotope dilution. ^{2}H, ^{13}C and ^{15}N tagged organic compounds for the isotopic dilution step will have to be synthesized where these are not now available. Under optimum conditions, accuracies of the order of ±1% are obtainable, but the more usual accuracy at present is in the ±5% range. (b) Gas chromatograph coupled to a mass spectrometer holds promise as a tool for the qualitative identification (or determination of purity) of organic constituents. (c) Radioisotope dilution using ^{3}H and ^{14}C tagged compounds, followed by spectral analysis and counting is capable of ±1% accuracy. (d) Calorimetric methods including differential scanning calorimetry and microcalorimetry are distinct possibilities, but depend upon isolation of the organic entity under test, or alternatively, the use of highly specific reactions (e.g. enzyme based) so that the measurement of the heat evolved can be attributed solely to the species under examination. (e) NMR for conformational and configurational evidence of separated species may also play a role.

5.1.2 *SRM's (pure substances)*. The primary source of SRM's for use in clinical chemistry is the US-NBS. Table 4 shows current availability.

How NBS-SRM's are characterized and certified, properties measured, techniques used, etc., have been previously reported in the literature.[19-21] Several other national laboratories are either planning or considering the issuance of clinical SRM's.

SRM's currently under development at NBS are shown in Table 5.

Obviously, many more SRM's for enzymes, steroids, and proteins are still needed and a concerted worldwide effort must be mounted if these needs are to be met on a reasonable time scale.

Table 4. Standard reference materials currently available from NBS

SRM	Type	Purity (%) or property
911a	Cholesterol	99·4
912	Urea	99·7
913	Uric acid	99·7
914	Creatinine	99·8
915	Calcium carbonate	99·9
916	Bilirubin	99·0
917	D-Glucose	99·9
918	Potassium chloride	99·9
919	Sodium chloride	99·9
920	D-Mannitol	99·8
921	Cortisol	98·9
922	Tris(hydroxymethyl)aminomethane, pH	99·9
923	Tris(hydroxymethyl)aminomethane, hydrochloride, pH	99·7
924	Lithium carbonate	100·0
925	VMA (4-hydroxy-3-methoxymandelic acid)	99·4
928	Lead nitrate	In preparation
929	Magnesium gluconate	In preparation
930b	Glass filters for spectrophotometry	Optical transmittance and absorbance
931a	Liquid filters for spectrophotometry	Optical transmittance and absorbance
932	Quartz cuvette for spectrophotometry	Optical transmittance
933	Clinical laboratory thermometers	0 and 25, 30, or 37°C
934	Clinical laboratory thermometer	0, 25, 30 and 37°C

5.2 *Reference methods and matrix reference materials*

The first reference method in clinical chemistry was that for calcium in serum. The basic philosophy underlying the development, the methodology used, the roles of the experts group, the laboratory network, and coordinating laboratory have been reported by Cali, Bowers and Young[22] and in greater detail by Cali and coworkers.[18]

Current reference method development efforts are being reported by Büttner, Jackson, Mitchell and Schaffer at this Symposium and reference to their papers is indicated to bring the reader up-to-date on these developments.

5.2.1 *Matrix reference materials.* A matrix reference material is, as the name implies, a reference material whose matrix is the same or similar to that of the material under analysis. One or more of the analytes in this matrix will be well-characterized (known with accuracy). This class of reference materials is, in principle, no different from the so-called "Q.C. sera materials," "serum reference materials," and others offered by manufacturers. The sine qua non for this classification scheme is that they should all be accurately characterized via a reference method—or, in initial development stages even by the definitive method if this should prove economically, or practically feasible. If the entire standardization process outlined in this paper were, in fact, in place and fully operational, the only difference between a matrix reference material issued, say, by NBS and a Q.C. reference sera issued by a manufacturer would be the degree of accuracy applied to each. Given the complexity of a large manufacturing process, the large amount of serum to be handled, etc. it does not seem reasonable, at least at this time, to expect the same degree of accuracy as might be anticipated from a small lot of material issued by a national laboratory.

Table 5. Under development (1975) list of clinical chemistry

SRM's (NBS) Name	Remarks and status
Bovine serum albumin	Research and development underway.
Liquid SRM's for spectrophotometry	To be certified at 240 and 300 nm; underway.
Potassium dichromate and potassium hydrogen phthalate	Dry form certified for absorbance at 240, 300, 400, 500, and 600 nm.
Cyanmethaemoglobin	International standard to be tested by NBS; (not to be issued by NBS—report only).
NADH	Research and development underway.
Sodium pyruvate	Research and development underway.
Quinine sulfate (powder)	Purity; corrected spectra; quantum efficiency. Research and development underway.[2]
Ionic activity SRM's	"Tris"* and phosphate in isotonic saline; calcium chloride, anhydrous; potassium fluoride (solid); pCO and pO_2-gases at 2 levels, underway.
Toxicology SRM's	99+% purity for phenobarbital, pentobarbital, and secobarbital diphenylhydantoin; research and development underway.

*(2-Amino-2-(hydroxymethyl)-1,3-propanediol.)

All one can say with certainty at this point is that matrix reference materials based on reference methodology are urgently needed, so that the transfer of accuracy into the local laboratory level can occur.

5.3 *Routine methods and secondary reference materials*

These methods and reference materials if based on the mechanisms proposed will provide for the medical accuracy for diagnosis and treatment that directly benefits the patient, and it is in these areas that the greatest impact will be felt. We take note of the fact that there is no paucity of routine methods or secondary reference materials (Q.C. sera, reference sera, etc.). In fact, the field is almost inundated with them and making rational choices among and between methods, reagents, kits, has become in itself a time-consuming activity. What is now required is the establishment of mechanisms for the evaluation of the accuracy of the routine methods and for the assurance of the well-characterized nature of the secondary reference materials. How this can be done in principle is now clear, what is not so evident is which organizations shall take the responsibility for getting on with the task, and how these activities shall be structured.

It is beyond the scope of this paper to lay out in detail which organizations should be involved and what structures should be built, but a few general remarks are appropriate.

International activities, by their very nature, should be limited to the coordination of actions by member bodies, the setting of priorities on a worldwide basis, and the gathering and dissemination of information. When and if the actual standards are under consideration, then the appropriate international standards body can help assure rational and harmonized standards.

Professional societies can and should play a leading role, especially at the national level, in coordinating, guiding and implementing many of the activities outlined. For example, the American Association of Clinical Chemists has now underway the development of reference methods for glucose and uric acid. These and similar efforts elsewhere need to be accelerated and intensified. In addition, the professional societies need to explore their potential for examining in a systematic way the accuracy—or rather the inaccuracy—of the multitudinous routine methods now extant.

In many countries these activities will be guided by or complemented by governmental agencies and laboratories. Significant increases in funding are required for this monumental task, and appropriate educational efforts aimed toward this end would seem to be in order. Regulatory functions by governmental agencies should be founded on the best scientific basis possible, and the accurate measurement system proposed and, indeed, now in process of development lends itself ideally to these activities.

The manufacturer's role in the production and characterization of secondary reference materials is crucial, for the system will only work in practice if there is an adequate transfer or maintenance of accuracy through the use of these materials. It would seem essential that every clinical chemist support and encourage those manufacturers that subscribe to and make every effort to build accuracy into his product through the mechanisms outlined. National standards bodies representing as they do both producer and consumer are also vital links in this process for it is in this forum that the practical realities are faced and resolved.

6. CONCLUSION

The theoretical basis for initiating measurement compatibility in clinical chemistry through accuracy is now in place. Through a hierarchy of methods and materials, traceability from the local laboratory level to the international base and derived units of SI can be achieved. Accuracy at the local level is assessed and maintained through the use of manufactured secondary reference materials whose accuracy, in turn, is based on reference methods of demonstrated accuracy and well-characterized matrix reference materials. In turn, reference methods and matrix reference materials are based on definitive methods and pure reference materials that are related to the SI units by direct experimental evidence of the highest quality.

It is the translation from this theoretical basis to practice that must now be addressed by scientific community of clinical chemists. Existing organizational structures should be utilized, rather than attempting to create new ones designed specifically for this task. The structure is complex and highly interactive. From the local laboratory to the international organization level, the structure is complex and highly interactive, and some existing organizations may need to adapt themselves internally in order to achieve the degree of cooperation and coordination required to build this edifice of accurate measurement.

In a cooperative spirit and with dedication on the part of all scientists in the field of clinical chemistry the task, while imposing, can be accomplished for the ultimate benefit of all mankind.

REFERENCES

[1]W. A. Shewart, Dept. of Agriculture, Graduate School, Washington, D.C. (1939).
[2]Metric Practice Guide, American Society for Testing and Materials, E-380-70, Philadelphia, PA (1970).
[3]ISO Recommendation, R-1000, reprinted by American National Standards Inst., New York (1969).
[4]Calibration and Test Services of the NBS, NBS Spec. Publ. 250 (1970).
[5]The National Standard Reference Data System. NBS Tech. Note 747 (1972).
[6]Time and Frequency. NBS Monograph 140 (1973).
[7]N. E. Dorsey and C. Eisenhart, *Sci. Monthly* **77**, 103–09 (1953).
[8]J. P. Cali and W. P. Reed, *Symposium on Trace Analysis.*
[9]H. H. Ku, Editor, NBS Spec. Publ. 300, Vol. 1 (1969).
[10]P. J. Campion, J. E. Burns and A. Williams, National Physical Laboratory, Her Majesty's Stationery Office, SBN 11 4800375, London, U.K. (1973).
[11]C. Eisenhart, *J. Res. NBS*, **67C**, 161 (1963).
[12]J. P. Cali, *Bull. Wld. Hlth. Org.* **48**, 721 (1973).
[13]J. P. Cali, *Med. Inst.* **8**, 17 (1974).
[14]R. N. Barnett, *Med. Inst.* **8**, 14 (1972).
[15]K. L. Churney and G. T. Armstrong, *J. Res. NBS* **72A**, (*Phys. and Chem.*), No. 5, 453 (1968).
[16]L. J. Moore and L. A. Machlan, *Anal. Chem.* **44**, 2291 (1972).
[17]L. J. Moore, Private communication (1975).
[18]J. P. Cali *et al.*, NBS Spec. Publ. 260–36 (1972).
[19]W. W. Meinke, *Anal. Chem.* **43**, 284 (1971).
[20]J. P. Cali *et al.*, NBS Monograph 148 (1974).
[21]R. W. Seward, Editor, NBS Spec. Publ. 419 (1975).
[22]J. P. Cali, G. N. Bowers, Jr. and D. S. Young, *Clin. Chem.* **19**, 1208 (1973).