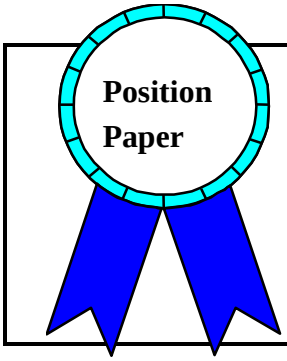
European
Diagnostic
Manufacturers
Association

Place Saint Lambert 14
1200 Woluwé St Lambert
Brussels, Belgium
+32 2 772 2225 tel
+32 2 772 2329 fax
edma@edma-ivd.eu
www.edma-ivd.eu



General aspects of Medical and Metrological Traceability in Laboratory Medicine

14 March 2001

INTRODUCTION

The In-vitro Diagnostic (IVD) Medical Device Directive of the EU (Directive 98/79/EC) requires in Annex I.A.3 that: "The traceability of values assigned to calibrators and/or control materials must be assured through available reference measurement procedures and/or available reference materials of higher order." In order to support the implementation of this requirement the European Commission has mandated two CEN/ISO Standards, prEN ISO 17511 on "Metrological traceability of values assigned to calibrators and control materials" and prEN ISO 18153 on "Metrological traceability of values for catalytic concentration of enzymes assigned to calibrators and control materials". These Standards explicitly refer to "metrological" traceability and to "calibrators" and "control material". Metrology is defined as the science of measurement. EDMA has prepared a specific document in order to provide IVD medical device manufacturers with some guidance on how to interpret these two upcoming Standards[1]. The CEN/ISO Standards exclude discussion of all items that are not strictly "metrological". However it has to be considered that meeting metrological requirements is not the primary goal of laboratory medicine, it is only a tool to serve the primary goal, which is generating useful and reliable information for medical decision-making.

In this connection many aspects have to be considered which are not mentioned in the two Standards. This document is intended to provide additional information about aspects of medical and metrological traceability in Laboratory Medicine which should be considered when applying the two CEN / ISO Standards. The document does not cover all aspects of traceability in Laboratory Medicine. Fields which are not covered by the two standards are not discussed (e.g. the morphological identification of cells, bacteria etc.)

1. Traceability in laboratory medicine

A primary goal of laboratory medicine is to contribute to optimal health care by generating reliable analytical results on patient samples, providing useful information to assist clinicians/physicians in *medical decision-making.* The interpretation of laboratory data is usually based on the comparison of an actual analytical result on a patient sample with decision-making criteria derived from prior clinical studies (and to some extent also from the prior experience of the clinician making the decision). Thus, a prerequisite for making correct medical decisions is that the actual pre-analytical, analytical and post-analytical conditions are consistent with the pre-analytical, analytical and post-analytical conditions underlying the decision-making criteria. Many factors are involved such as patient condition and medical history, method of collecting and handling patient samples, analytical procedure and mode of expressing and transmitting results to clinicians. We call this *medical traceability*

[1] EDMA Guidance Document "Interpretation of the CEN/ISO Standards prEN ISO 17511 and prEN ISO 18153 on metrological traceability of values assigned to calibrators and control materials", January 2001

of laboratory results. Appropriate medical decisions are only possible if this *medical traceability* is ensured.

2. Medical and metrological traceability

The consistency of the analytical process is an essential element of the *medical traceability*. This means that a calibrator used to standardize a routine method must be traceable to the calibrator and method used in the clinical studies upon which the decision-making criteria are based. Historically, there are very few clinical studies where reference measurement procedures of higher metrological order were directly employed. In most clinical studies the routine measurement procedures used were not made traceable to reference measurement procedures and reference materials of higher metrological order, mostly due to the fact that the reference system elements (reference materials and/or reference measurement procedures) were simply not available. Thus, many of the established decision-making criteria for assessing laboratory data currently employed in medical practice are based solely on data obtained with routine measurement procedures. This holds especially true for immunoassays where it usually happens that the first assay for a particular quantity, which had been clinically evaluated, was the one which established the numerical database for future medical decision-making. Once these initial numerical values are reported in the literature and become clinically accepted, other manufacturers interested in placing an assay method for the same quantity onto the market, frequently will adjust the calibration of their assays in order to report values consistent with the first assay (at least in the clinically relevant range).

When tracing calibration of a routine measurement system back to reference materials and reference measurement procedures of higher metrological order (metrological traceability) any change in the trueness of measurement has to be evaluated. If there are significant changes in patient results, the *medical traceability* has to be re-established by introducing new decision-making criteria.

For control materials intended to verify the trueness of routine measurement procedures, it is important that their target values are consistent with the trueness underlying the medical decision-making criteria applied. Assigning target values with reference measurement procedures of higher metrological order only makes sense if the routine measurement procedure can be traced back to that reference measurement procedure, and the control materials in question are commutable (i.e. they demonstrate performance with the routine measurement procedure which is consistent with the performance of patient samples).

3. Analytical reference systems

3.1 General remarks

Tracing back the calibration of routine measurement procedures to metrologically higher order reference materials and reference methods does not automatically result in comparability of the results of different routine measurement procedures. Comparability of results between methods can only be achieved if a well-designed reference system is available to which routine measurement procedures can be traced. Such reference systems consist of four essential elements, including:

- a clear biochemical definition of the analyte with regard to the intended clinical use
- a reference measurement procedure which specifically measures the analyte (or quantity) in human samples

- suitable reference materials (primary and secondary matrix-based)
- information regarding the *medical traceability* of the analytical reference system.

3.2 Defining and classifying the analytes

A detailed definition of the quantity to be measured is an indispensable part of an analytical reference system in laboratory medicine. This is often neglected since the analyte definition is frequently taken for granted, through abbreviation to the simplistic name of the analyte alone (e.g. sodium, prostate specific antigen etc.). This approach creates problems if the analyte is not a uniform substance but represents a group of substances. For example, often quite different forms of the analyte are simultaneously present, and form-specific methods (e.g. HCG methods designed for pregnancy detection and methods designed for diagnosing/monitoring tumours) are intended to differentiate these particular forms by measuring specifically one form and not the other. Therefore, in approaching the harmonization of laboratory tests by defining their traceability back to reference systems of higher metrological order, it is necessary to define two groups of analytes. These are:

Type A quantities - physico-chemically well defined compounds. Their concentrations can be expressed in molar units (SI unit). Examples are electrolytes, glucose, cholesterol, urea, uric acid, steroid hormones etc.

Type B quantities - quantities that do not represent a uniform substance but are typically a heterogeneous mixture of substances. For these quantities, it is usually unknown which post-translationally modified, degraded or complexed forms are present in a particular human fluid sample. Frequently, there are differences among individuals, and even within an individual the composition of a given analyte (quantity) can vary depending on time and/or physico-pathological situation at the moment the sample is collected. The analyte in these cases is a heterogeneous mixture of many different molecular species that are measured as “the” analyte, using one or more chemical properties which all species of the analyte have in common (e.g. common epitopes in immunoassays). The concentrations of Type B quantities are expressed in terms of arbitrary units (e.g. WHO International Units or sometimes in “artificial” molar units by referring to a reference preparation).
The group of Type B quantities is much larger than that of Type A and currently includes an estimated 500 analytes. Examples of Type B analytes are tumour markers and parameters for coagulation, endocrinology, immunology, virology, and bacteriology. For Type B quantities the definition of the analyte should also refer to the intended clinical application, since this indicates which species of the analyte should be measured preferentially (e.g. determination of specific ferritin species in human serum for diagnosing anaemia or for monitoring tumours.)

As all classification of natural phenomena is a simplification, there is, of course, also some heterogeneity in Type A quantities, which should be considered (e.g. free and protein-bound calcium, sodium ion concentration or sodium ion activity; various forms of bilirubin; steroid hormone binding to specific binding proteins, etc.). However, the complexity of Type A quantities is generally considered to be much less than that of Type B analytes.

3.3. Reference materials

There is a significant difference in available reference materials for Type A and for Type B quantities. For Type A quantities, reference materials containing the analyte as a pure compound (primary

reference material) can usually be prepared. This is much more complicated for Type B quantities. Since these analytes are heterogeneous and their composition in human body fluids varies, all reference materials for Type B quantities are by definition only **surrogates** for the analytes to be measured in patient samples. While such materials may resemble to some extent the typical heterogeneous mixture of the analyte present in the human fluids, it often represents only an 'average' condition. Therefore, primary reference materials for type B quantities do not exist. In lieu of primary reference materials, EDMA believes that the most appropriate surrogates are **panels of human samples,** collected from both normal individuals and patients with defined pathology, carefully selected to reflect the clinical application(s) of the analyte. The use of such panels by manufacturers is considered by some metrological purists to be a questionable practice. These metrologists frequently express a preference for artificially prepared calibrators, due to their well-defined and reproducible nature. Unfortunately, most artificial materials do not adequately represent the molecular structure of the analyte as found in the human samples where they are measured, due to lack of a complete specification for the diverse molecular forms and the complex matrix present in human samples. Therefore, EDMA recommends that artificial reference materials should only be used if and when their performance has been proven, adequately validated, and shown to ensure the *medical traceability* of the laboratory test.

For a number of Type B analytes WHO International Standards currently exist. However, it cannot be taken for granted that all these materials are suitable reference materials to support metrological traceability of routine measurement procedures in laboratory medicine. The purified biological substance contained in these materials is often a heterogeneous mixture itself (e.g. due to variations in degree of glycosylation), and bears a resemblance to the analyte to be measured in the human fluid only in the sense of biological activity. These materials are only suitable if they are appropriate surrogates of the analyte to be measured in human samples with regard to the intended clinical application, and therefore should be used by manufacturers with considerable caution.

3.4 Reference measurement procedures

For Type A analytes reference measurement procedures which specifically measure the analyte and are independent of the routine measurement procedures can be developed, and for many of these analytes such procedures are already available. The amount of the analyte as such or in a matrix can be accurately measured, and the values assigned to calibrators or control materials are generally method-independent "true values" (within an uncertainty interval). If various reference measurement procedures are available for a given analyte, all the reference measurement procedures should yield the same results (within uncertainty limits) for the amount of analyte in a given sample. It should not matter which reference procedure is used for assigning a value to a reference material or calibrator.

In contrast, for Type B quantities, reference measurement procedures independent of routinely employed measurement procedures are at present lacking in virtually all cases. This situation is well known, but has been overlooked in the interest of pragmatism. In fact, in EQA schemes for Type B quantities measured by routine immunoassays *overall* mean values of results are often reported and called "consensus values" or "reference

values". These "consensus" or "reference" values are then used incorrectly to calculate systematic error (bias) for particular IVD products and individual laboratories.

A way to achieve harmonization of routine measurement procedures for Type B quantities is via the development of *International Conventional Reference Measurement Procedures and International Conventional Reference Materials*. To accomplish this, firstly the analyte intended to be measured should be defined in analytical terms taking into account the clinical application of the analyte. Secondly, a reference measurement procedure independent of routine measurement procedures should be devised. Thirdly, a reference material, preferably in the matrix to be measured routinely, should be defined and established. Since the IVD industry serves a global market, these three points must be addressed by a **global** consortium of stakeholders, including international professional scientific and medical organizations, metrology institutes and the IVD industry. Such a process has already been initiated for establishing international reference measurement systems for a number of enzymes.

3.5 Information regarding the medical traceability of an analytical reference system

With regard to the priority of *medical traceability*, currently related (usually) only to routine measurement procedures, method comparison studies between the reference measurement procedure and the major routine measurement procedures using human samples should be performed. In this way it can be understood what impact on medical decision-making the implementation of a new reference system would have. Points to consider in evaluating these data include (1) the impact on laboratory and medical practice, (2) potential consequences regarding the medical decision-making criteria, and (3) how to safeguard the legacy of historical clinical experience if a new reference system leads to a change in reference intervals or other medical decision-making criteria. These considerations are of great importance before deciding to undertake the introduction of a new reference system in laboratory practice. EDMA believes that these issues have not been adequately taken into account in the past, and that only metrological issues were considered of primary importance.

4. Establishing metrological traceability of routine measurement procedures to analytical reference systems

It is an important goal in laboratory medicine to achieve *comparability* of results among different routine measurement procedures so that the results may become interchangeable over time and space. Comparability of data on a global scale would contribute to improvements in health care since results of clinical studies undertaken in different locations or times could be applied universally. A way to overcome the current diversity in results for routine methods is to establish traceability of results for all routine measurement procedures to reference systems (reference materials and reference measurement procedures) of higher metrological order. EDMA supports this concept, provided that such analytical reference systems are agreed upon by consensus and internationally accepted, and the technical points discussed in the present paper are carefully considered. With regard to *medical traceability*, any change in trueness of routine measurement procedures resulting from calibration adjustments that might be needed to implement traceability to a reference system should be carefully evaluated. If

there is a change, the decision-making system for the routine method must be adjusted accordingly, and clinicians have to be educated concerning the change. In order to maintain the value of clinical experience, correlation of measurement results obtained with the new calibration to results of measurements obtained with the previous calibration should be established. Adjusting the decision-making criteria is of outstanding importance since, even if from a metrological point of view the routine analytical method is biased, clinicians can still arrive at correct clinical decisions if the decision-making criteria they apply incorporate the same bias. In contrast, they could arrive at incorrect clinical decisions if patient results are true with regard to the reference system, but the decision-making criteria are only valid by using the previous calibration for the test.

In recent years, a number of efforts have been initiated to harmonize measurement results for Type B quantities which are often routinely measured with various immunoassays. EDMA believes that the optimal approach to this problem is to identify those components of the substance of clinical interest in the human fluid intended to be measured, and to establish and validate measurement procedures to assign values to human samples and candidate calibrators.

An excellent example of this approach is the recent IFCC project for standardization of measurement of HbA_{1c}. For this project, glycated haemoglobins were identified in blood of diabetic patients, then clinically relevant components of the various molecular forms were selected, and finally a reference measurement procedure was developed, i.e. a combination of high-pressure liquid chromatography and electron-spray mass spectrometry[2]. In the opinion of EDMA, this example represents a good model for establishing an international conventional reference system.

To promote the concept of an "international conventional reference system", EDMA believes that it is the responsibility of international scientific and medical bodies and representatives of the IVD industry, working in collaboration, to review the state-of-the-art in measurement systems for a prioritised list of Type B quantities, and to offer recommendations for potential solutions which are scientifically valid and clinically useful. The priority of quantities for this review should be governed by their clinical importance. In order to avoid confusion and waste of resources, EDMA opposes the development and establishment of any national or regional analytical reference systems for Type B quantities.

5. The concept of uncertainty

For decades the analytical reliability of test results for particular measurement procedures in laboratory medicine has been characterised in terms of random error as estimated by the precision (imprecision) and systematic errors related to the accuracy (now called trueness) of the method. Some years ago a new concept was launched adopted from statistical techniques applied in industrial settings– the concept of **uncertainty** of measurement. This concept is inclusive of all components of variation that could influence the estimate of the value indicated (e.g. random and systematic components, including components due to equipment variation, environmental factors, etc.). Methodology for determining and expressing the

[2] Hoelzel,W; Miedema K . Development of a reference system for the international standardization of HbA1c/glycohemoglobin determinations. JIFCC 1996; 9: 62-7.

Finke A, Kobold U, Hoelzel W, Weykamp C, Miedema K, Jeppson J-O . Preparation of a candidate primary reference material for the international standardization of HbA1c determinations. Clin Chem Lab Med 1998; 36: 299-308, and Miedema K. Electrospray mass spectrometry for measurement of glycohaemoglobins (editorial). Clin Chem 1997; 43:705-707

"uncertainty of measurement" is described in a document known as the "GUM", published by ISO in 1993[3]. The GUM was established by a consortium of metrology experts, mainly analytical chemists and physicists. Since the GUM is difficult to interpret, a group of analytical chemists, Eurachem Europe, produced a practical guide to the application of the GUM in the field of chemistry in general[4]. According to the GUM, uncertainty of measurement must be determined across the entire chain of traceability, starting with the provider of the reference material, extending through to the IVD reagent manufacturer and his processes for assignment of values to calibrators supplied with reagent kits, and ultimately to the end user of the test system, to the final test result reported to the clinician. This concept is new in laboratory medicine, and its full implementation is expected to take some years and require significant educational efforts. There is also controversy among some experts who prefer existing concepts of measurement error, because of the confusion that might be generated, especially among laboratory users, by introducing the new concept.

EDMA's position is that neither the question of whether or not the concept of measurement uncertainty should be introduced, nor the responsibility for the significant educational efforts required for the implementation of the new concept, should be the burden of IVD manufacturers (e.g. in instructions for use, on product labels etc.). Rather, laboratory medicine scientific and medical professional societies should be responsible for resolving all the issues related, and undertaking any associated educational efforts. Therefore, EDMA advises IVD manufacturers to implement, in-house, the determination of the uncertainty of measurement of the values assigned to their product calibrators and trueness control materials, and to maintain these data in their product-related Technical Documentation, as required in prEN ISO 17511 (clause 6). However, it is not required that manufacturers provide any values for measurement uncertainty in instructions for use or on labels. Professional users may, however, obtain this information upon request to manufacturers.

6. Medical traceability, metrological traceability , uncertainty and analytical performance goals

Medical traceability, metrological traceability, uncertainty and analytical performance goals are strongly related. Metrological traceability becomes meaningless if the uncertainty of the value assigned to the various calibrators in the traceability chain confounds medical decision-making. When a certain level of uncertainty is exceeded, the *medical traceability* is no longer guaranteed even if formally an unbroken traceability chain from a reference material of higher order to patient values can be established. This is of great practical significance for Type B quantities since the values of the currently available international reference materials often have a high uncertainty. This is due to the fact that, in the absence of a reliable reference measurement procedure, the assigned values are consensus values generated by taking an average of results from different routine measurement procedures (e.g. WHO standard for TSH[5]). EDMA believes that introducing the concept of uncertainty requires performance goals for allowable uncertainties, which should be based on medical needs. Such goals do not currently exist. All existing published

[3] Guide to the expression of uncertainty of measurement, ISO 1993

[4] Quantifying uncertainty in analytical measurement; Eurachem, 1st edition 1995, 2nd edition 2000.

[5] Rafferty B and Gaines Das R. Comparison of pituitary and recombinant human thyroid-stimulating hormone (rhTSH) in a multicenter collaborative study: establishment of the first World Health Organization Reference Reagent for rhTSH. Clin Chem. 1999; 45: 2207-2215

papers on analytical performance goals refer to the concepts of precision and accuracy (now called trueness[6,7]) and indicate allowable values for analytical imprecision (typically expressed as CVs) and systematic deviation (bias).
A major concern is that the medical meaning of uncertainty is still rather unclear. Metrologists recommend that all uncertainties generated across the entire traceability chain, beginning with the hierarchically highest-placed calibrator, should be imposed on patient results and reported to the clinician. It is the opinion of EDMA that doing this would be a mistake and would result only in confusion for clinicians. This is because these uncertainty figures, in some cases, may be very large. Since all the current clinical information and decision-making criteria have been generated with routine measurement procedures, the average metrological uncertainty is already accounted for in current medical decision-making practice. Presenting total uncertainty figures to clinicians would only result in a loss of credibility for many laboratory tests. Therefore, the introduction of the concept of metrological uncertainty of laboratory results first requires a careful re-evaluation of the medical decision-making process when using laboratory data.

References

Documents, publications or other material are referenced in footnotes in the text of this document.

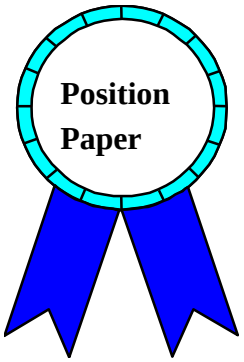
Authors

Wieland G. Hoelzel
Rudolf M. Lequin
Rob Slobbe
Gian Alfredo Scassellati

Acknowledgement

The comments from Dr. Neil Greenberg, member of the ISO/TC 212/WG2, are most gratefully acknowledged

EDMA Standardization Working Party March 2001



[6] EDMA Guidance Document "Specific Performance Characteristics of IVDMDs, October 2000
[7] ISO 5725-1: 1994 - Accuracy (trueness and precision) of measurement methods and results.