Minireview

Standardization in clinical enzymology: a challenge for the theory of metrological traceability

Ilenia Infusino¹, Gerhard Schumann², Ferruccio Ceriotti³ and Mauro Panteghini¹*

¹ Enzyme Reference Laboratory, Center for Metrological Traceability in Laboratory Medicine (CIRME), University of Milan, Milan, Italy
² Institute for Clinical Chemistry, Medical School Hannover, Hannover, Germany
³ Diagnostica e Ricerca San Raffaele S.p.A., IRCCS San Raffaele, Milan, Italy

Abstract

The goal of standardization for measurement of the catalytic concentration of enzymes is to achieve comparable results in human samples, independent of the reagent kits, instruments, and laboratory where the assay is performed. To pursue this objective, the IFCC has established reference systems for the most important clinical enzymes. These systems are based on the following requirements: a) reference methods, well described and evaluated extensively; b) suitable reference materials; and c) reference laboratories operating in a highly controlled manner. When these reference systems are used appropriately, the diagnostic industry can assign traceable values to commercial calibrators. Clinical laboratories that use procedures with validated calibrators to measure human specimens can now obtain values that are traceable to higher-order reference procedures. These reference systems constitute the structure of the traceability chain to which the routine methods can be linked via an appropriate calibration process, provided that they have a comparable specificity (i.e., they are measuring the same catalytic quantity).

Keywords: calibration; enzymes; IFCC reference methods; reference values; traceability.

Introduction

Enzymatic measurements are important for the diagnosis and monitoring of diseases of the liver, pancreas, skeletal muscle, bone, etc. (1–3). This emphasizes the importance of standardized measurement results in practice. However, variability of the results from different laboratories can be observed (4). If enzyme results are not comparable, clinical laboratories should adopt method-dependent reference intervals. However, laboratories often use reference intervals that are different without any valid reason. This is difficult to explain; possible reasons could be the adoption of data from the literature or the manufacturer's recommendations, without any critical appraisal. Also, this can be related to scarce attention paid to this issue such that changes in analytical methodologies are not accompanied by corresponding changes in reference intervals. A typical example of the situation can be derived from a survey on reference intervals used in Italy in 2005 for alanine aminotransferase (ALT). In more than 60 laboratories, all purporting to use the same method (even though on different analytical platforms), the upper reference limit (URL) for males ranged from 40 to 72 U/L, while the lower limit ranged from 0 to 30 U/L (5). In addition to the lack of comparability of results, this not uncommon situation is dangerous and confounding both for clinicians and patients since the same analytical result can be considered "normal" in one laboratory and "abnormal" in another one, according to the reference interval being used.

Achieving interlaboratory agreement of enzyme activity measurements that may permit the use of common reference intervals represents one of the most important standardization efforts in Laboratory Medicine. The goal of this article is to provide an update on the present knowledge in the field of standardization of enzyme measurements since our previous review (6), with specific reference to new advancements made since publication of the previous summary.

Suggested approaches to achieve enzyme standardization

In the past, one way to improve the comparability of enzyme results was to promote the widespread use of optimized analytical methods (7). However, the approach of “method globalization” for enzyme standardization has shown insurmountable limitations. Thus, the goal of a single, universal method to measure the catalytic concentration of a given enzyme in daily practice has not been achieved (8).

Current efforts to standardize enzyme measurements should allow comparability of results for human serum samples, independent of the test kits and instruments used. To achieve this goal, the “reference system” approach, based
on the concepts of metrological traceability and a hierarchy of measurement procedures, has recently been proposed (9–11). In applying the reference system theory, enzymes represent a special class of analytes (12). They are defined in terms of the ‘‘catalytic amount’’, which is the amount of an agreed-upon substrate that is converted to product in an agreed-upon measurement system. The numerical results of catalytic activity measurements depend entirely on the experimental conditions under which measurements are made. Therefore, a reference measurement procedure, which defines the conditions under which the activity of a given enzyme is measured, occupies the highest level of the traceability chain (13). In addition to the reference measurement procedure, the system also requires reference materials for the intermediate transfer of trueness from the reference procedure to the routine laboratory assays. Clinical laboratories that use commercial methods with validated calibrators to measure patient samples may finally obtain values that are traceable to the reference measurement procedure. Also, the obtained values are independent of the particular method or instrument, finally permitting comparability of enzyme results (Figure 1).

However, the applicability of the enzyme reference system concept is possible only if the reference materials used to transfer trueness to the field methods are commutable. Commutability in enzymology has been defined as ‘‘the ability of an enzyme (reference or control) material to show inter-assay activity changes similar to those of the same enzyme in human serum’’ (14). If commutable reference materials suitable for direct calibration of field methods are lacking, a panel of native human sera, with values certified by the reference measurement procedure and acting as a secondary reference material, represents the only possible alternative for establishing traceability to the reference system (6). Calibration of the commercial system must be in accordance with correlation results obtained using the value-assigned human samples (15).

The implementation of standardization in enzymology through a reference system requires that commercial methods have similar analytical specificities toward the specific enzyme when compared to the reference procedure. For example, it will not be possible to definitively align procedures for transaminases that do not incorporate pyridoxal-5’-phosphate to a procedure that does, such as the IFCC reference procedure. This is because the ratio of holoenzyme to apoenzyme differs between human serum samples.

**Tools for enzyme standardization**

To specifically promote the establishment of reference measurement systems (RMS) in clinical enzymology, the IFCC created a Working Group in 1997 that was later upgraded to a Committee (16).

**Primary reference measurement procedures**

The first objective of the group was to select suitable primary reference measurement procedures to underpin the accuracy of the system and traceability of measurement results. A decision was made to modify, when existing, the original IFCC methods recommended during the 1980s for the measurement of catalytic concentrations of enzymes at 30°C, avoiding, however, sample blanking and changing the reaction temperature to 37°C, with a thorough re-evaluation of incubation times and linearity (17). Measurement conditions should be described in the form of a detailed standard operating procedure (SOP). Also, in order to achieve low-levels of measurement uncertainty, careful control of all metrological aspects related to gravimetry, volume, pH, reaction temperature, and photometry has to be promoted (Table 1). Thus, the resulting uncertainty for all relevant steps of the analytical procedures can be known permitting all major components of the uncertainty budget to be kept under control (Figure 2).

IFCC reference methods for the measurement of creatine kinase (CK), lactate dehydrogenase (LDH), ALT, aspartate aminotransferase (AST), γ-glutamyltransferase (GGT), and α-amylase (AMY) are available. Also, a candidate reference procedure for alkaline phosphatase (ALP) is ready for approval and publication (18–24). The IFCC Committee is also working on a concept for the development of a reference method for pancreatic lipase (LPS), favoring spectrophotometry as the measurement principle (25).

**Accredited reference laboratories**

From the beginning, reference procedures have been validated and tested for transferability in a network of reference laboratories.

**Table 1**  Aspects to be carefully controlled in performing reference measurement procedures for enzymes.

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Control Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gravimetry</td>
<td>Controlled by calibrated test weights</td>
</tr>
<tr>
<td>Volumetry</td>
<td>Controlled by gravimetry</td>
</tr>
<tr>
<td>Temperature</td>
<td>Controlled by calibrated thermometer</td>
</tr>
<tr>
<td>pH</td>
<td>Controlled by calibrated equipment</td>
</tr>
<tr>
<td>Photometric wavelength</td>
<td>Controlled by certified filters or solutions of holmium</td>
</tr>
<tr>
<td>Photometric absorbance</td>
<td>Checked by test solutions certified by a national metrology institute</td>
</tr>
</tbody>
</table>
Figure 2: Overview of relevant uncertainty components of the enzyme measurements using reference procedures.

Reference materials

After the first step of enzyme standardization consisting of the development of primary reference measurement procedures and establishment of reference laboratories, the improvement of consistency of results may be achieved using appropriate reference materials for trueness transfer to a lower hierarchical level. The IFCC and the Institute for Reference Materials and Measurements (IRMM) have cooperated to certify reference materials for GGT, LDH, ALT, CK, AMY and AST (Table 2) (27, 28). As mentioned previously, the commutability of the materials intended to be used for calibration of commercial systems is a key issue and the main criterion required for the transfer of trueness to routine enzyme methods. For the reported reference materials, commutability has only been shown for a restricted number of methods, and additional information is still needed before using them to ensure comparability of results for certain methodologies, for instance procedures using dry chemistry (6). In addition, these monoenzyme reference materials are available in limited amounts, are relatively expensive to purchase, and therefore, are not routinely used to calibrate enzyme assays. Thus, additional multienzyme materials that behave in a manner that is similar to human samples would be very useful (29, 30).

Common reference intervals

An additional issue associated with standardization efforts is the need to develop scientifically sound and globally useful reference intervals for serum enzyme catalytic concentrations. Lack of proper reference intervals may indeed hamper the implementation of standardization in enzymology.
Table 2  Characteristics of the enzyme reference materials certified by the IFCC in cooperation with the Institute for Reference Materials and Measurements (IRMM).

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Code</th>
<th>Origin</th>
<th>Form</th>
<th>Certified concentration, U/L</th>
<th>Uncertainty, U/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGT</td>
<td>ERM-AD452</td>
<td>Pig kidney</td>
<td>Light subunit</td>
<td>114.1</td>
<td>± 2.4</td>
</tr>
<tr>
<td>LDH</td>
<td>ERM-AD453</td>
<td>Human erythrocytes</td>
<td>LDH₁ isoenzyme</td>
<td>502.0</td>
<td>± 7.0</td>
</tr>
<tr>
<td>ALT</td>
<td>ERM-AD454</td>
<td>Pig heart</td>
<td>–</td>
<td>186.0</td>
<td>± 4.0</td>
</tr>
<tr>
<td>CK</td>
<td>ERM-AD455</td>
<td>Human heart</td>
<td>MB isoenzyme</td>
<td>101.0</td>
<td>± 4.0</td>
</tr>
<tr>
<td>AMY</td>
<td>IRMM/IFCC 456</td>
<td>Human pancreas</td>
<td>Pancreatic isoenzyme</td>
<td>546.0</td>
<td>± 18.0</td>
</tr>
<tr>
<td>AST</td>
<td>ERM-AD457</td>
<td>Recombinant</td>
<td>Liver cytosolic isoenzyme</td>
<td>104.6</td>
<td>± 2.7</td>
</tr>
</tbody>
</table>

As: a) the implementation of standardization can modify enzyme results, b) without adequate reference intervals, this situation can impair the interpretation of the results and, paradoxically, worsen the patient’s outcome, c) the absence of reliable reference intervals for newly standardized commercial methods may hamper their adoption, and d) usually, a single clinical laboratory or manufacturer may not have the means to adequately produce reference intervals.

The RMS represents a trueness-based approach. With this approach, different commercial methods that provide results traceable to the system are able to produce comparable results in clinical laboratories using these assays. Thus, reference intervals obtained with analytical procedures that produce results traceable to the corresponding RMS can be transferred between laboratories (becoming “common”), providing that they use commercial assays that produce results traceable to the same reference system and populations have the same characteristics, or, alternatively whether it is known that the specific enzyme is not influenced by ethnicity or environmental factors (Figure 4) (31). The definition of common reference intervals should hopefully cause the disappearance of different intervals employed for the same enzyme, providing more effective information to clinicians.

Using the approach described above, some preliminary examples of common reference intervals for enzymes can already be found in the literature. In Caucasian subjects, the reference interval for CK was found to be 46–171 U/L for males, and 34–145 U/L for females, when measured with an assay traceable to the IFCC 37°C reference procedure (32). In another study, the reference interval for GGT activity in adult men was 11–49 U/L when data were produced using three assays traceable to the IFCC reference procedure (33). Finally, the reference interval for LDH activity in adult Caucasian subjects, determined at 37°C using a procedure traceable to the IFCC reference method, was found to be 125–220 U/L (34). Recently, AST, ALT, LDH and GGT pediatric reference intervals have been developed using assays with traceability to the IFCC reference procedures at 37°C (35).

Large multicenter studies are needed for a robust definition of common reference intervals, using a protocol for collaborative experiments that include well defined prerequisites (Table 3). The difficulties are related to the need for verifying traceability of participating laboratories by the distribution of commutable frozen sera, with values assigned by the reference measurement procedure. Other difficulties include the co-ordination among participating centers for the performance of thousands of tests and enrolment of hundreds of individuals, which entails considerable cost. Particularly, in the development of reference intervals, the methods that are employed must produce results that are traceable to the RMS for that specific enzyme. For this reason, the trueness of participating laboratories should be verified and, if necessary, experimental results corrected in accordance with correlation results with the reference procedure. Alternatively, the samples from reference individuals can be collected at the different centers, frozen and then shipped to a central laboratory where all the enzymatic analyses are performed. The latter approach is simpler and allows better control of the analytical phase. However, this approach uses frozen samples and thus introduces a variable not typical for the clinical laboratories, and raise some doubts for some unstable enzymes such as ALT.

In summary, the production of common reference intervals for enzymes may pose numerous practical problems to solve.

Figure 4  Common reference intervals as fourth pillar of the reference measurement system: how a problem becomes a solution.

Table 3 Prerequisites for the organization of multicenter studies for defining common reference intervals.

- Only clinical laboratories using methods for which the traceability to primary reference procedures and/or materials is clearly stated by the manufactures are eligible for participation
- Only clinical laboratories strictly following the manufacturer’s instructions are eligible for participation
- Goals for maximal allowable interlaboratory variability, i.e., maximum allowable bias, must be stated ‘a priori’ and verified in a preliminary experiment
- The analytical quality during the production of reference values must be carefully monitored by an ‘ad hoc’ quality control program.
However, the possibility of providing reference intervals that are applicable to any laboratory, able to produce results traceable to the reference measurement procedure, seems to be quite realistic (36).

**Status of enzyme traceability implementation**

In the European Union, the implementation of result traceability in Laboratory Medicine to available RMS is mandatory by law (37). However, the introduction of correctly standardized assays in enzymology is a complicated task. Overall, it appears that in many cases method bias can be reduced by better calibration to the internationally accepted reference systems. Even for commercial assays using methodological principles that differ in analytical specificity when compared with the internationally recommended reference systems, these assays should be replaced by analytical procedures where the traceability of calibration to the corresponding IFCC reference measurement procedure has been experimentally proved (38).

In 2006, a study involving 70 European laboratories assessed enzyme assays from six major manufacturers for traceability to IFCC RMS. This study used a commutable serum-based material targeted with ALT, AST, CK, GGT, LDH and AMY reference procedures (39). Results from commercial methods were assessed using a system where the maximum allowable error was derived from the desirable analytical performance based on the biological variation model. Of the enzyme measurements, CK and ALT results were very good. For AST and GGT, only two company systems would fully comply. Measurements of LDH and AMY still had major drawbacks, suggesting the need for major improvement. This was primarily the result of using methods with different analytical specificities for these enzymes, leading to results that were not traceable to the internationally accepted reference system. In some cases, diagnostic companies may still prefer to produce enzyme methods with non-IFCC traceable calibrations. This will allow laboratories to choose between different marketed assays and/or calibrations, some of which are clearly not traceable to the RMS.

**Post-market surveillance: the way forward**

Although a more correct implementation of the reference system concept by the manufacturers is required, it is the responsibility of our profession to verify the accuracy and comparability of the commercially available enzyme methods (Table 4). A major role must be accomplished by the External Quality Assessment Scheme (EQAS) organizers through the use of commutable materials with values assigned by IFCC reference procedures (Table 5). True value assignment to EQAS materials allows objective evaluation of the performance of enzyme measurements through an accuracy-based (instead of inferior consensus-based) grading of the competency of participating clinical laboratories, by applying clinically allowable total error limits derived from biological variability data (Table 6).

<table>
<thead>
<tr>
<th>Feature</th>
<th>Aim</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Values assigned with IFCC reference procedures by an accredited reference laboratory</td>
<td>To check trueness as traceability to IFCC reference systems</td>
</tr>
<tr>
<td>• Proved commutability of control material(s)</td>
<td>To allow transferability of results to patient samples</td>
</tr>
<tr>
<td>• Definition of the clinically allowable total error of measurements</td>
<td>To permit reliable application of laboratory measurements in clinical setting</td>
</tr>
</tbody>
</table>

**Table 4** Major steps in the achievement of standardization of enzyme measurements.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Aim</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Values assigned with IFCC reference procedures by an accredited reference laboratory</td>
<td>To check trueness as traceability to IFCC reference systems</td>
</tr>
<tr>
<td>• Proved commutability of control material(s)</td>
<td>To allow transferability of results to patient samples</td>
</tr>
<tr>
<td>• Definition of the clinically allowable total error of measurements</td>
<td>To permit reliable application of laboratory measurements in clinical setting</td>
</tr>
</tbody>
</table>

**Table 5** Main features for the applicability of enzyme true value concept in External Quality Assessment Schemes.

<table>
<thead>
<tr>
<th>Quality level</th>
<th>Optimum, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>AST</td>
<td>± 15.2</td>
</tr>
<tr>
<td>ALT</td>
<td>± 32.1</td>
</tr>
<tr>
<td>GGT</td>
<td>± 22.2</td>
</tr>
<tr>
<td>LDH</td>
<td>± 11.4</td>
</tr>
<tr>
<td>CK</td>
<td>± 30.3</td>
</tr>
<tr>
<td>AMY</td>
<td>± 14.6</td>
</tr>
</tbody>
</table>

Total error goals were calculated as: bias goal + (1.65 × imprecision goal). Bias and imprecision goals (desirable and optimum) were derived from intrasample and intersample biological variabilities (available at http://www.westgard.com/biodatabase1.htm) of the respective enzymes, according to (40).

In conclusion, the establishment of traceability to enzyme reference systems provides the clinical laboratory and medical community with a universal means of ensuring result comparability. This can be obtained without disruptive changes to existing working methods, or to an individual laboratory’s preference for an analytical system.

**Conflict of interest statement**

Authors’ conflict of interest disclosure: The authors stated that there are no conflicts of interest regarding the publication of this article.

Research funding: None declared.

Employment or leadership: None declared.

Honorarium: None declared.
References


