

SESSIONE 3

STANDARIZZAZIONE IN ENZIMOLOGIA CLINICA: LA TEORIA DELLA RIFERIBILITA' METROLOGICA ALLA VERIFICA PRATICA

Sala F

Mercoledì 12 Ottobre 2005, ore 10.15 - 12.45

S3.1

IFCC REFERENCE SYSTEMS FOR ENZYME STANDARDIZATION

Panteghini M.

Cattedra di Biochimica Clinica e Biologia Molecolare Clinica, Dipartimento di Scienze Cliniche "Luigi Sacco", Facoltà di Medicina e Chirurgia - Polo di Vialba, Università degli Studi di Milano, Via G.B. Grassi 74, 20157 Milano, Italy

The primary goal of standardization for measurements of catalytic concentrations of enzymes is to achieve comparable results in human samples, independent of the reagent kits instruments and laboratory where the procedure is carried out. In order to pursue this objective, the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) has established reference systems for the most important clinical enzymes. These systems are based on three hinges: a) reference measurement procedures extensively evaluated and carefully described, b) certified reference materials (RMs), and c) a network of reference laboratories operating in a highly controlled manner. The original IFCC-recommended procedures for alanine aminotransferase (ALT), aspartate aminotransferase (AST), creatine kinase (CK), g-glutamyltransferase (GGT), lactate dehydrogenase (LDH), a-amylase (AMY) and alkaline phosphatase (ALP) were modified to optimize them at 37°C with the definition of detailed operating procedures. A small group of laboratories, world-wide located perform these procedures manually, with self-made reagents on carefully calibrated instrumentations. In cooperation with the Institute for Reference Materials and Measurements (IRMM), five RMs, that were already available, have been re-certified by these laboratories for ALT, CK, GGT, LDH and AMY activities. In addition, recombinant human AST and ALP materials have been selected as candidate RMs and a certification campaign will begin very soon. Using these RMs and the manufacturer's standing procedures, industry can assign traceable values to commercial calibrators. Clinical laboratories, which will use routine procedures with these validated calibrators to measure

human specimens, may finally obtain values which are traceable to reference procedures. In conclusion these reference systems constitutes the structure of the traceability chain to which the routine procedures can be linked via an appropriate calibration process, provided that they have a comparable analytical specificity (i.e. they are measuring the same quantity).

S3.2

STATE OF STANDARDIZATION IN ENZYMOMETRY; RESULTS FROM AN EUROPEAN PILOT PROJECT

Carlo Franzini

Università degli Studi di Milano, Milano (Italy)

Main aim of this international study was to assess the state-of-the art in the measurement of 7 clinically useful serum enzyme activities with particular reference to the traceability of results to the IFCC reference method values. The study design included the distribution of 1-mL aliquots of deep-frozen special serum preparation to 5 laboratories in 3 Countries (Germany, Italy and The Netherlands), each using one of 6 previously chosen, largely available analytical systems. Each participant was asked to assay the serum in 5 replicates using his/her routine analytical system. The study included some special features. First, reference laboratories assigned the values of catalytic activity concentration (7 enzymes: ALT; AST, CK; LD; GGT; AMY; ALP) to the special serum with IFCC reference methods (with the exclusion of ALP, whose value was assigned by consensus). Second, the commutability of the test serum with patient sera was verified for a number of available analytical systems. In this way, measurement trueness could be evaluated by comparison with reference methods results, and the results were transferable to patient sera assay in "routine" condition. This report concerns essentially the analysis of the results obtained in our Country (Italy). Possible detection and elimination of outlier (or incongruous) results was a main problem in data evaluation. Although this operation may be always considered to be arbitrary, it was thought that data could

be discarded when it appeared or it was known that: ALT and AST were assayed without P-5-P in the mixture; LD was measured in the Pyr@Lact instead of in the Lac@Pyr direction; DEA instead of AMP was the buffer in ALP assay. Following such criteria a total of 18 series of results (90 single results) were eliminated, considering a series the set of 5 replicate measurements performed in one laboratory. Due to the number of laboratories really participating, and to the elimination of outliers, the total number of "Italian" useful values was 835, a figure that still compares reasonably with the planned number (1050), and allows statistically robust conclusions. The within laboratory imprecision was from excellent to good: out of 42 mean CV% values, 37 were <1.5%, 4 were in the interval from 1.5% to 2.0%, 1 reached a peak value of 4.6%. The inter-lab/within system imprecision was higher: out of 42 CV% values 25 were <3.0%, 12 spanned the interval from 3.0% to 5.0%, 5 were >5% (peak value: 7.8%).

The overall imprecision values were still higher, spanning the interval from 3.9% to 10.0%. The trueness was assessed by comparing the measured to the expected (reference or consensus) values calculating the bias as: $\text{bias\%} = (\text{measured} - \text{expected}) / \text{expected} \times 100$. With the exception of amylase, the within-system bias values for each enzyme were, alternatively, negative or positive spanning the interval from -19.0% to +14.0%; for amylase the bias values were in the interval from -40.0% to -2.4%. Due to mutual compensation of negative and positive values, the overall bias values for 6 of the 7 enzymes were in the interval from -3.0% to +2.8%, for amylase the overall bias value was -16.0%. Concerning the inter-countries comparison, overall CV and mean values were made available from the project's organizer, Dr. Rob T. P. Jansen (The Netherlands). The number of useful results was: Germany, 833; Italy, 886; The Netherlands 460. Results from this elaboration are not fully comparable with our elaboration of the Italian results, because of different criteria in the elimination of incongruous results, as above outlined. However Country-specific results are each other comparable because of the application of homogeneous criteria. Country specific overall CV values (from 4.9% to 10.8%) showed to be substantially comparable, with the main exceptions of lactate dehydrogenase, where CV values as high as 43.5% and 48.1% were found, and amylase, with CV values of 13.0% and 19.3%, possibly as the result of averaging results obtained under different standardizations of measurement. This is also reflected in the overall bias values. Out of 21 bias values, 16 spanned the interval from -6.7% to +1.8%, but values as high as +22% and -32% were observed for lactate dehydrogenase and amylase, respectively. We can conclude that: 1) even within the same analytical system a few routine enzyme measurements are not traceable to the recommended IFCC standardization; 2) the high intrinsic repeatability of the analytical systems is not fully taken advantage of, possibly because of defects of calibration. The results of

this project show that there is room for improvement and in which direction to address future efforts.

S3.3

HOMOGENEITY OF REFERENCE INTERVALS AND DECISION LEVELS IN CLINICAL ENZYMOLOGY: THE FINAL TARGET HAS BEEN MET?

Ferruccio Ceriotti

Diagnostics e Ricerca S. Raffaele S.P.A. Ospedale San Raffaele, Via Olgettina 60, 20132 Milano, Italy

The goal of the IVD Directive, requiring traceability to higher metrological levels, is to allow the production of comparable results independently from the method used. Due to the standardization efforts of IFCC and JCTLM this goal could now be reached also in the field of clinical enzymology that, until now, was the reign of method dependent results (by definition catalytic activity depends upon the method or, better, is a property of the analytical systems itself). The definition of a reference system for the measurement of the concentration of catalytic activity of the most common enzymes and the publication of the list of reference measurement and materials on the JCTLM web site is driving towards the homogeneity the results, but what about reference ranges?

If we look at the present situation the picture is disconcerting: on a database of about 590 Italian laboratories for ALT measurement more than 30 different upper reference limits are indicated just for males, spanning from 29 to 83 U/L. Forty two % of the laboratories claims identical intervals for males and females. If we restrict the field to the methods that claim to be "according to IFCC with pyridoxal phosphate" (only 47 labs), we still find 13 different upper reference limits just for males (from 31 to 70 U/L) with lower limits spanning from 0 to 30 U/L (lower and upper limits almost overlap!). The confusion of different analytical principles combined to laboratory dependent reference intervals creates a dangerous mix. And substituting the decision levels for the reference ranges, instead of normalizing the situation adds confusion to confusion. In fact usually, due to the difficulty of defining a unique decision level, the multiples of upper reference limits (URL) are adopted. It can be easily demonstrated that the combination of different methods with inappropriate reference ranges reduces the comparability of the clinical information provided instead of increasing it. To try to improve this unacceptable situation and to eventually complete the effort of standardization produced by the JCTLM and IFCC, IFCC itself has started a committee with the task of defining appropriate reference intervals for the most relevant measurements, including enzymes.

These reference intervals will be obtained through a collaborative effort of an international network of laboratories and will probably be adopted by the

manufacturers that provide test kits traceable to the IFCC reference methods. The adoption of these common reference ranges should be the first step on which to develop appropriate decision levels for the most common clinical situations.

S3.4

THE IVD INDUSTRY - AN IMPORTANT PARTNER IN THE PROCESS OF ENZYME STANDARDISATION

Hinzmann R., Sysmex Europe GmbH, Norderstedt, Germany

Enzyme concentrations are usually not measured as mass concentrations but as catalytic concentrations, depending on the temperature, pH, buffer, ionic strength presence of co-enzymes, activators, substrate, etc. The peculiarity of these tests is that the analyte is defined by the measurement conditions. Therefore, contrary to the reference methods for substrates such as creatinine or glucose the reference methods for enzymes are consensus-based. Standardization is in the interest of the IVD industry and has been pursued for many years, but most often at a national level, leading to non-comparable results and requiring manufacturers to provide different kits for different countries. The IVD industry supports IFCC's international enzyme standardization approach.

Questions in this context are:

(1) Which enzymes? To justify the financial engagement wide-spread use of the test is mandatory (cost / effort relationship).

(2) Which methods?

Manufacturers must be able to develop traceable routine methods or to adjust existing ones: Practical laboratory requirements, however, often mean that routine methods differ significantly from the reference procedures defining the analyte causing problems with control materials and sometimes also patient samples. Examples are shorter pre-incubation and serum start instead of substrate start to increase the sample throughput. Sometimes liquid, ready to use reagents with a long shelf-life and calibration stability with some components added or changed are used.

(3) Acceptance of the clinical community?

It is important that the standardised - and thereby changed - analyte still reflects clinical requirements. And, unfortunately, standardization of methods has not yet resulted in harmonization of reference intervals even for equivalent reference populations. This is also required to make the standardization clinically useful.

(4) How to standardize?

A good approach is to use a set of patient samples to trace the routine method back to the reference procedure

in order to avoid commutability issues.

(5) Which reference laboratory to choose and how to organize the logistics?

Manufacturer's own laboratory or laboratory from IFCC reference laboratory network. (Long-term availability of the reference laboratory?) Cost considerations.

These questions need to be discussed between IFCC and manufacturers before the standardization begins.

S3.5

REFERENCE INTERVALS OF ENZYMES: THE STATE OF THE ART FOCUSED THROUGH AN EXTERNAL QUALITY ASSESSMENT SCHEME (EQAS)

Secchiero S., Sciacovelli L., Zardo L., Plebani M. Centro di Ricerca Biomedica - Castelfranco Veneto (TV)

EQAS aim at improving the performance of laboratories by way of education, scientific recommendations and standardization taking into account clinical needs and quality specifications.

The participants to the EQAS managed by the Centre of Biomedical Research (CRB) submit their results on a medical report form with the indication of the Reference Intervals (RI) adopted, and, in order to stimulate the use of appropriate RI we provide diagrams showing each laboratory the result obtained on a control sample against its own RI and with those of other Participants. In 2004 we included the study of some clinical cases within the EQAS for Clinical Biochemistry on serum in order to stimulate participants to provide not only the analytical results but also a clinical interpretation and suggestions for additional laboratory investigations.

One of the clinical cases studied concerned a female 35 years old, affected by a cholestatic syndrome recognizable by means of the results of the hepatic functionality indexes: Total Bilirubin = 26.8 mmol/L, Direct Bilirubin = 15.0 mmol/L, AST = 64 U/L, ALT = 61 U/L, GGT = 282 U/L. ALP (AMP) = 400 U/L, LDH (Piruvate à Lactate) = 776 U/L and ALP isoenzymes = presence of a biliary fraction further to the normal bone-liver isoenzyme.

We studied the RI of the 450 Laboratories that participated to this clinical case: 170 directly afferent to CRB and 280 to the EQA group of S. Orsola-Malpighi of Bologna with which we co-operate. The study pointed out a large variability in RI of hepatic enzymes, for example among the 170 CRB Participants 16, 10, 31, 16 and 10 gave for AST, ALT, GGT, ALP and LDH, respectively, RI detaching significantly from the other using the same diagnostic system.

The medians of the rate result/Max RI value of each laboratory showed that some diagnostic systems

presented a significant statistically difference from the others. For example, for AST, the medians of the rates of Abbott (1.56), Beckman (1.53) and IL (1.58) were lower than Roche,Hit/Mod (1.92) and Dade-Behring (1.94); for ALT, the medians of the rates of Dade-Behring (1.25), Abbott (1.31) and Ortho (1.35) were lower than Roche,Integra (1.71) and Olympus (1.64); for GGT, the medians of the rates of Ortho (4.66) and Beckman (5.35) were lower than Roche,Integra (8.75), IL (8.30) and Olympus (7.93).

The reason of these discrepancies probably comes out from the fact that the laboratories often adopt the RI issued on the package insert of diagnostic system utilized which are often inadequate to the results provided by the diagnostic system itself. In conclusion, a lot of work remains to do to standardize the IR of enzymes and information obtained from EQAS are useful both for the Laboratories and for the manufacturers.