Clin Chem Lab Med 2012;50(7):1237–1241 © 2012 by Walter de Gruyter • Berlin • Boston. DOI 10.1515/CCLM.2011.791

Opinion Paper

Implementation of standardization in clinical practice: not always an easy task

Mauro Panteghini*

Centre for Metrological Traceability in Laboratory Medicine (CIRME), University of Milan, Milan, Italy

Abstract

As soon as a new reference measurement system is adopted, clinical validation of correctly calibrated commercial methods should take place. Tracing back the calibration of routine assays to a reference system can actually modify the relation of analyte results to existing reference intervals and decision limits and this may invalidate some of the clinical decision-making criteria currently used. To maintain the accumulated clinical experience, the quantitative relationship to the previous calibration system should be established and, if necessary, the clinical decision-making criteria should be adjusted accordingly. The implementation of standardization should take place in a concerted action of laboratorians, manufacturers, external quality assessment scheme organizers and clinicians. Dedicated meetings with manufacturers should be organized to discuss the process of assay recalibration and studies should be performed to obtain convincing evidence that the standardization works, improving result comparability. Another important issue relates to the surveillance of the performance of standardized assays through the organization of appropriate analytical internal and external quality controls. Last but not least, uncertainty of measurement that fits for this purpose must be defined across the entire traceability chain, starting with the available reference materials, extending through the manufacturers and their processes for assignment of calibrator values and ultimately to the final result reported to clinicians by laboratories.

Keywords: quality assessment; reference material; reference measurement procedure; result comparability; standardization; traceability.

*Corresponding author: Professor Mauro Panteghini, Laboratorio Analisi Chimico Cliniche, Ospedale Luigi Sacco, Via GB Grassi 74, 20157 Milano, Italy
Phone: +39 02 3904 2806, Fax: +39 02 503 19835,
E-mail: mauro.panteghini@unimi.it
Received July 27, 2011; accepted October 26, 2011;
previously published online February 29, 2012

Introduction

The issue of standardization in Laboratory Medicine represents an absolute priority for public health (1). Quoting Bossuyt et al. (2), "standardization of laboratory tests has an ethical dimension as it aims to affect the way diagnostic tests are performed in order to guarantee optimal care for patients in a global world". Sure enough, laboratory customers (i.e., doctors and patients) "simply" expect test results to be accurate and comparable and interpreted in a reliable and consistent manner (3). There is now an international agreement on the fact that, to be accurate and comparable, results must be traceable to high-order references (4, 5). The main objective of traceability approach is to enable the results obtained by the calibrated routine procedure to be expressed in terms of the values obtained at the highest available level of the calibration hierarchy (6, 7). In addition to be interpreted in a reliable and consistent manner, results should be compared with an appropriate population reference interval (transversal evaluation at biological level), an appropriate decision limit (transversal evaluation at nosological level) or with a previous result from the same individual (longitudinal evaluation) (8). It is therefore clear that the standardization of patients' results through the availability and the practical application of a working reference measurement system (which includes reference materials, reference procedures and accredited laboratories performing these procedures) is not enough to fulfill expectations of clinicians and patients if a proper way of interpretation of these standardized results is not simultaneously available.

Need to establish "clinical" traceability

There are now several examples showing that the adoption of the metrological approach to standardize assay results in clinical laboratories can substantially modify the relation of analyte results to existing reference intervals and decision limits (9–12). Without adequate reference intervals and/or decision limits, this situation can impair the interpretation of the results and, paradoxically, worsen the patient's outcome. Consequently, the absence of reliable reference intervals and/or decision limits for the newly standardized commercial methods may significantly hamper their adoption in clinical practice (13, 14).

The knowledge about the clinical validity of laboratory tests and the decision-making criteria used by physicians are often

based on data that are method-dependent, generated with routine tests that are not standardized. As reported before, tracing back the calibration of routine tests to a reference measurement system (i.e., implementing standardization) can modify the analyte results and this may invalidate some of the clinical decision-making criteria currently used (4). Thus, to maintain the accumulated clinical experience, the quantitative relationship to the previous calibration system should be established and, if necessary, the clinical decision-making criteria should be adjusted accordingly to ensure that patient care remains consistent despite the changes (7). In addition, if needed, establishing traceability should also improve specificity of test results and lead to a qualitative improvement of its clinical significance (15). Some practical examples can clarify this issue.

Example 1: serum creatinine and glomerular filtration rate equations

A reference measurement system for standardization of measurements of creatinine in serum is now available and virtually all major global manufacturers have recalibrated their creatinine assays to isotope dilution-mass spectrometry (IDMS) reference method worldwide (15–17). Since creatinine results, which were used to generate the clinical validation for the equations often used to estimate kidney function, such as Modification of Diet in Renal Disease (MDRD) study, Cockcroft-Gault or Schwartz formulas, were not traceable to the reference measurement system, tracing back the calibration of routine creatinine tests to this system may invalidate the clinical value of glomerular filtration rate (GFR) equations originally proposed (18). In order to avoid this risk, a re-expression of these equations with standardized creatinine results has, therefore, been advocated. As a matter of fact, a MDRD equation, sometimes referred to as "175" formula, was re-expressed for GFR calculation (eGFR) with IDMS standardized serum creatinine results with the best approximation (19). By using this equation and standardized creatinine assays, clinical laboratories can report eGFR more uniformly and accurately (20). On the contrary, there is no modified Cockcroft-Gault equation available for use with the IDMS-traceable creatinine results and the risk of over-estimating eGFR using this formula at the present time is very high (9, 21). Consequently, its use to estimate kidney function as the basis for drug dose adjustment recommendations should be discontinued (22). Finally, for eGFR in pediatrics a revised IDMS-traceable Schwartz equation has recently been reported (23). As for pediatric patients the specificity of creatinine methods remains a major concern, this updated formula is, however, only valid in combination with enzymatic methods for creatinine (9, 24).

Example 2: prostate-specific antigen

It is now well-demonstrated that recalibrating a prostate-specific antigen (PSA) assay from an original "Hybritech" calibration to the new WHO International Reference Preparation (IRP) 96/670 in order to assure assay traceability to this higher-order material results in approximately 20%–25% lower PSA values (10, 25). Consequently, a PSA cut-off of 3.0–3.1 μg/L should be considered for WHO calibrated assays in order to achieve the same diagnostic efficiency as with a cut-off of 4.0 μg/L in traditionally calibrated assays (26). Studies have shown that the application of the WHO standard for PSA assays with traditionally used PSA thresholds leads to a significant decrease in detection of prostate cancer (27, 28). For example, applying the WHO calibration for screening a cohort of 5865 men yielded a 19% decrease in prostate biopsies and a 20% decrease in detected cancers compared with the "Hybritech" calibration, at a cut-off for biopsy of 4.0 μg/L (28). On the other hand, a recent study has shown that the standardization of PSA assays is still limited because only a buffer-based reference material (i.e., the WHO IRP) has been prepared and because PSA does not have a total reference measurement system including commutable reference materials and a reference measurement procedure (29). To further improve PSA standardization, a reference measurement system should be established to transmit the trueness of the primary reference material to patient samples, which is the final goal (30, 31).

Example 3: glycated hemoglobin

Also for glycated hemoglobin (HbA_{1c}) the knowledge about the clinical validity of the test and the corresponding decision-making criteria used by physicians are substantially based on data which are method-dependent (32, 33). This required the appraisal of the existing relationship between results traceable to the IFCC reference measurement system for HbA_{1c} and those from methods used in landmark clinical studies, once the reference system became available (34). Indeed, a reliable and stable correlation has been demonstrated, allowing the conversion of analytical and clinical data from one system to another through the use of the so-called "master equation" (35). Consequently, it has been possible to derive adequate target (decision) limits for the clinical use of HbA_{1c}, when measured with methods traceable to the reference measurement system and expressed using the recommended SI unit, namely "mmol/mol" (11). HbA_{1c} is also a good example on how to practically solve problems related to the implementation of standardization when recalibration is not only a matter of a modified slope of the regression line, but can also be a matter of an intercept, showing that the determined quantity by use of some routine procedures is not the same as that defined by the reference system.

Need to define the clinically acceptable limits for validation of metrologically traceable calibrations

Establishing traceability of results for a test measurement should be inseparably linked to the definition of acceptable measurement uncertainty to fit the intended clinical application ("fitness for purpose") (36). Some years ago, Thienpont et al. (37) already pointed out that the absence

of clearly defined tolerable deviations derived from clinical needs "might results in a large gray zone with respect to the extent of traceability expected from diagnostic manufacturers, partially or totally invalidating its theoretical advantages". Today, there is a substantial agreement that the uncertainty of measurement that fits for purpose must be defined across the entire traceability chain, starting with the provider of reference materials, extending through the diagnostic manufacturers and their processes for assignment of calibrator values, and ultimately to the final result reported to clinicians by end users (i.e., the clinical laboratories) (38, 39). This approach should be applied to every analyte measured in the clinical laboratory in order to establish if the current status of the uncertainty budget of its measurement associated with the proposed metrological traceability chain is suitable for clinical application of the test (40). Using the measurement of albumin in serum as a model, we recently demonstrated that the reference measurement system (and the associated uncertainty) currently available for this protein is probably not enough to guarantee the accuracy needed for its clinical usefulness (Figure 1) (44). Particularly, we should probably recognize that the uncertainty associated to each step of the traceability chain should be reduced to obtain a final combined standard uncertainty of the patient's sample result that may fulfill quality levels for albumin measurement established using, e.g., the information on the biological variation of the analyte (44).

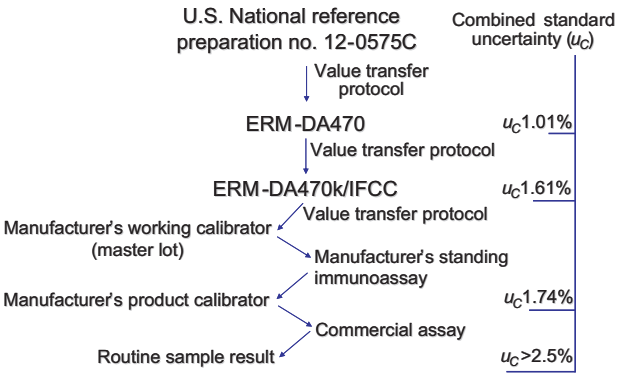

Figure 1 Metrological traceability chain for albumin measurement in serum and associated combined standard uncertainties.
For serum albumin, the availability in the past of the reference preparation ERM-DA470 (Institute for Reference Materials and Measurements) and more recently of the ERM-DA 470k/IFCC Human Serum Proteins reference material, both including albumin in the list of certified plasma proteins, together with immunochemical methods based on turbidimetry – nephelometry principles, recognized as reference measurement procedures by the Joint Committee on Traceability in Laboratory Medicine, provides the basis to maintain the assay traceability to the US National Reference Preparation no. 12-0575C, representing the highest level of the albumin traceability chain. The standard uncertainties of ERM-DA470, ERM-DA470k/IFCC and manufacturer's product calibrator used to calculate the combined standard uncertainty at different levels of the traceability chain were obtained from Refs. (41–43), respectively. For more details, see ref. (44).

How the uncertainty goal-setting should occur is the object of an ongoing debate. Using the consensus established at the IFCC-IUPAC Stockholm Conference for setting quality specifications in Laboratory Medicine (45), the best scientific approach of defining analytical performance goals should rely on (in a hierarchical order): 1) the evaluation of the effect of analytical performance on clinical outcome in the specific clinical setting, 2) data based on components of biological variation and, if the previous information is lacking, 3) data based on clinical and laboratory experts' opinion and published recommendations. As the outcome-based information is missing for the majority of analytes, the approach using biological variation data has generated more attention, even if recent publications have questioned the reliability and robustness of available information (46).

Role and duties of diagnostic manufacturers

The European Union (EU) Directive on In Vitro Diagnostic (IVD) medical devices obliges diagnostic manufacturers to ensure traceability of their analytical systems to recognized higher-order references (47). Thus, the primary onus is on the manufacturers to drive traceability. Manufacturers are responsible for implementing suitable commercial systems that fulfill this requirement and for documenting the measurement uncertainty that fits for purpose (36). Individual clinical laboratories should therefore rely on the manufacturers who must ensure traceability of their analytical systems to the highest available level. It should be noted that, according to the current concepts focusing on requirements and responsibilities of IVD manufacturers, the "analytical system" components offered by a given supplier should include platform, reagents, calibrators and control materials for checking the correct system alignment (see below).

Given the importance of the manufacturers' role, they should be involved from the beginning in the standardization projects. Particularly, dedicated meetings with manufacturers should be organized to discuss changes, e.g., the process of assay recalibration, and the corresponding time schedule (48). Furthermore, a proof-of-concept study should properly be performed to obtain convincing evidence that the standardization works by improving result comparability in clinical setting, in order to justify costs required to recalibrate assays due to manufacturing process changes, communications to regulatory agencies, etc. (49).

Need to monitor the efficacy of traceability implementation on a continuous basis

According to IVD regulations manufacturers should provide test components including control materials that shall give the user confidence that the values obtained by use of the test system are traceable. Confidence means that the recoveries for the control materials are close to 100% and the

acceptance limits are as narrow as needed for the medical requirements. Clinical laboratories should verify the consistency of the performance declared by the manufacturers of their analytical systems during daily routine operations performed in accordance with the manufacturer's instructions. This requires that the internal quality control used to monitor the analytical performance of the employed system is organized into two independent components: the former related to analyze the system control material(s) and to confirm that current measurements are acceptable ("unbiased") according to the manufacturer's established parameters (i.e., the acceptable range of control materials) and the latter (using a different control material) designed to evaluate the system imprecision (36, 39).

External Quality Assessment (EQA) programs provide an assessment of laboratory performance and of analytical quality of measurements. However, to ensure that traceability is maintained some mandatory requirements must be fulfilled (14, 36, 39). First, EQA material values must be assigned with reference measurement procedures by an accredited reference laboratory. This will permit one to check the measurement uncertainty of participating laboratories against the established reference measurement systems. Second, EQA materials should display proved commutability to allow transferability of participating laboratory performance to patient samples. Finally, the clinically allowable uncertainty of measurements should be defined in order to verify the suitability of laboratory measurements in a clinical setting (36).

Conclusions

Accumulated experience is now showing that standardization projects not only have to address metrological traceability, but should also consider the efficacy of its implementation. Of course, the implementation of standardization in clinical practice needs first the availability of the three main pillars of reference measurement systems, i.e., reference measurement procedures, reference materials and accredited reference laboratories. Then, it needs to define a fourth pillar represented by traceable reference intervals/decision limits. Finally, an appropriately organized analytical (internal and external) quality control should become the fifth pillar. All of the stakeholders (clinical biochemists, IVD manufacturers, EQA organizers) need to co-operate to implement the standardization in a concerted action following a well-established time schedule. Last, but not least, implementation approaches should be developed in collaboration with clinicians and their associations to ensure that patient care remains consistent despite scheduled changes (50).

Acknowledgments

Presented in part at the IFCC-WorldLab Congress held in Berlin, Germany, May 2011; Symposium "Traceability, standardization, harmonization: concepts and implementation".

Conflict of interest statement

Author's conflict of interest disclosure: The author 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

1. Hernandez JS, Dale JC, Bennet KE, Varkey P. Challenges and opportunity for medical directors in pathology and laboratory medicine. Am J Clin Pathol 2010;133:8–13.
2. Bossuyt X, Louche C, Wiik A. Standardisation in clinical laboratory medicine: an ethical reflection. Ann Rheum Dis 2008;67:1061–3.
3. McLawhon RW. Patient safety and clinical effectiveness as imperatives for achieving harmonization inside and outside the clinical laboratory. Clin Chem 2011;57:936–8.
4. Panteghini M. Traceability as a unique tool to improve standardization in laboratory medicine. Clin Biochem 2009;42:236–40.
5. Vesper HW, Thienpont LM. Traceability in laboratory medicine. Clin Chem 2009;55:1067–75.
6. ISO 17511. In vitro diagnostic medical devices – measurement of quantities in biological samples – metrological traceability of values assigned to calibrators and control materials. Geneva, Switzerland: ISO, 2003.
7. White GH. Metrological traceability in clinical biochemistry. Ann Clin Biochem 2011;48:393–409.
8. Büttner J. Laboratory findings: structure, validity and significance for medical cognitive processes. Eur J Clin Chem Clin Biochem 1991;29:507–19.
9. Delanghe JR, Cobbaert C, Harmoinen A, Jansen R, Laitinen P, Panteghini M. Focusing on the clinical impact of standardization of creatinine measurements: a report by the EFCC Working Group on Creatinine Standardization. Clin Chem Lab Med 2011;49:977–82.
10. Stephan C, Kahrs AM, Klotzek S, Reiche J, Muller C, Lein M, et al. Toward metrological traceability in the determination of prostate-specific antigen (PSA): calibrating Beckman Coulter Hybritech Access PSA assays to WHO standards compared with the traditional Hybritech standards. Clin Chem Lab Med 2008;46:623–9.
11. Panteghini M, John WG. Implementation of haemoglobin A_{1c} results traceable to the IFCC reference system: the way forward. Clin Chem Lab Med 2007;45:942–4.
12. Thienpont LM, Van Uytfanghe K, Beastall G, Faix JD, Ieiri T, Miller WG, et al. for the IFCC Working Group on Standardization of Thyroid Function Tests. Report of the IFCC Working Group for Standardization of Thyroid Function Tests; Part 2: Free thyroxine and free triiodothyronine. Clin Chem 2010;56:912–20.
13. Ceriotti F, Hinzmann R, Panteghini M. Reference intervals: the way forward. Ann Clin Biochem 2009;46:8–17.
14. Infusino I, Schumann G, Ceriotti F, Panteghini M. Standardization in clinical enzymology: a challenge for the theory of metrological traceability. Clin Chem Lab Med 2010;48:301–7.
15. Panteghini M. Enzymatic assays for creatinine: time for action. Clin Chem Lab Med 2008;46:567–72.
16. Panteghini M, Myers GL, Miller GW, Greenberg N. The importance of metrological traceability on the validity of creatinine measurement as an index of renal function. Clin Chem Lab Med 2006;44:1187–92.

17. Miller WG. Estimating glomerular filtration rate. Clin Chem Lab Med 2009;47:1017–9.
18. Panteghini M. Estimating glomerular filtration rate from serum creatinine measurements: analytical issues and standardization programs. RIMeL/IJLaM 2007;3(Suppl):56–60.
19. Levey AS, Coresh J, Greene T, Marsh J, Stevens LA, Kusek JW, et al. Expressing the Modification of Diet in Renal Disease study equation for estimating glomerular filtration rate with standardized serum creatinine values. Clin Chem 2007;53:766–72.
20. Levey AS, Coresh J, Greene T, Stevens LA, Zhang Y, Hendriksen S, et al. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. Ann Intern Med 2006;145:247–54.
21. Stevens LA, Manzi J, Levey AS, Cheng J, Deysher AE, Greene T, et al. Impact of creatinine calibration on performance of GFR estimating equations in a pooled individual patient database. Am J Kidney Dis 2007;50:21–35.
22. Narva AS. Assessment of kidney function for drug dosing. Clin Chem 2009;55:1609–11.
23. Schwartz GJ, Muñoz A, Schneider MF, Mak RH, Kaskel F, Warady BA, et al. New equations to estimate GFR in children with CKD. J Am Soc Nephrol 2009:20:629–37.
24. Miller WG. Estimating equations for glomerular filtration rate in children: laboratory considerations. Clin Chem 2009;55:402–3.
25. Fillée C, Tombal B, Philippe M. Prostate cancer screening: clinical impact of WHO calibration of Beckman Coulter Access prostate-specific antigen assays. Clin Chem Lab Med 2010;48:285–8.
26. Stephan C, Köpke T, Semjonov A, Lein M, Deger S, Schrader M, et al. Discordant total and free prostate-specific antigen (PSA) assays: does calibration with WHO reference materials diminish the problem? Clin Chem Lab Med 2009;47:1325–31.
27. Link RE, Shariat SF, Nguyen CV, Farr A, Weinberg AD, Morton RA, et al. Variation in prostate specific antigen results from 2 different assay platforms: clinical impact on 2,304 patients undergoing prostate cancer screening. J Urol 2004;171:2234–8.
28. Jansen FH, Roobol M, Bangma VC, van Schaik RH. Clinical impact of new prostate-specific antigen WHO standardization on biopsy rates and cancer detection. Clin Chem 2008;54:1999–2006.
29. Stephan C, Siemßen K, Cammann H, Friedersdorff F, Deger S, Schrader M, et al. Between-method differences in prostate-specific antigen assays affect prostate cancer risk prediction by nomograms. Clin Chem 2011;57:995–1004.
30. Ishibashi M. Standardization of prostate-specific antigen (PSA) assays: can interchangeability of PSA measurements be improved? Clin Chem 2006;52:1–2.
31. Panteghini M. Standardization of cardiac troponin I measurements: the way forward? Clin Chem 2005;51:1594–7.
32. The Diabetes Control and Complications Trial (DCCT) Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med 1993;329:977–86.
33. UK Prospective Diabetes Study (UKPDS) Group. Intensive blood glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). Lancet 1998;352:837–53.
34. Hoelzel W, Weykamp C, Jeppsson JO, Miedema K, Barr JR, Goodall I, et al. IFCC reference system for measurement of hemoglobin A1c in human blood and the national standardization schemes in the United States, Japan, and Sweden: a method-comparison study. Clin Chem 2004;50:166–74.
35. Weykamp C, John WG, Mosca A, Hoshino T, Little R, Jeppsson JO, et al. The IFCC reference measurement system for HbA_{1c}: a 6-year progress report. Clin Chem 2008;54:240–8.
36. Panteghini M. Application of traceability concepts to analytical quality control may reconcile total error with uncertainty of measurement. Clin Chem Lab Med 2010;48:7–10.
37. Thienpont LM, Van Uytfanghe K, Rodriguez Cabaleiro D. Metrological traceability of calibration in the estimation and use of common medical decision-making criteria. Clin Chem Lab Med 2004;42:842–50.
38. Armbruster D. Measurement traceability and US IVD manufacturers: the impact of metrology. Accred Qual Assur 2009;14:393–8.
39. Panteghini M. Traceability in laboratory medicine: copernican revolution or activity for a restricted professional club? 4th International Scientific Meeting "Rethinking quality control in the traceability era". November 30, 2010, Milano, Italy. Available at: http://users.unimi.it/cirme/public/UploadAttach/Panteghini2010.pdf. (Accessed July 17, 2011).
40. Stöckl D, Sluss PM, Thienpont LM. Specifications for trueness and precision of a reference measurement system for serum/plasma 25-hydroxyvitamin D analysis. Clin Chim Acta 2009;408:8–13.
41. Whicher JT. BCR/IFCC reference material for plasma proteins (CRM 470). Community Bureau of Reference. International Federation of Clinical Chemistry. Clin Biochem 1998;31:459–65.
42. Zegers I, Keller T, Schreiber W, Sheldon J, Albertini R, Blirup-Jensen S, et al. Characterization of the new serum protein reference material ERM-DA470k/IFCC: value assignment by immunoassay. Clin Chem 2010;56:1880–8.
43. Traceability and Uncertainty Cobas c501/c311-C.f.a.s. PUC. Roche Diagnostics GmbH, Centralized Diagnostics Operations. Standardization and Quality Control, Penzberg. March 2009.
44. Infusino I, Braga F, Mozzi F, Valente C, Panteghini M. Is the accuracy of serum albumin measurements suitable for clinical application of the test? Clin Chim Acta 2011;412:791–2.
45. Hyltoft Petersen P, Fraser CG, Kallner A, Kenny D, eds. Strategies to set global analytical quality specifications in laboratory medicine. Scand J Clin Lab Invest 1999;59:475–585.
46. Braga F, Dolci A, Mosca A, Panteghini M. Biological variability of glycated hemoglobin. Clin Chim Acta 2010;411:1606–10.
47. Dati F. The new European Directive on in vitro diagnostics. Clin Chem Lab Med 2003;41:1289–98.
48. Panteghini M, John G, Mosca A, Weykamp C. Implementation of standardization of HbA_{1c} measurement. Summary of the meeting with manufacturers held in Milan, Italy, December 12, 2007. Clin Chem Lab Med 2008;46:573–4.
49. Xia C, Tong Q, Wang Q, Tang Z, Qi L, Chi S, et al. Application of five frozen human-pooled serum samples assigned by the International Federation of Clinical Chemistry and Laboratory Medicine reference procedure in a traceability investigation of gamma-glutamyltransferase catalytic concentration measurements in China. Ann Clin Biochem 2010;47:189–94.
50. Mosca A, Branca MT, Carta M, Genna ML, Giorda CB, Ghidelli R, et al. Recommendations for the implementation of international standardization of glycated hemoglobin in Italy. Clin Chem Lab Med 2010;48:623–6.