



Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/clinchim

A multicentric evaluation of IDMS-traceable creatinine enzymatic assays

Laurence Piéroni^a, Pierre Delanaye^{b,*}, Anne Boutten^c, Anne-Sophie Bargnoux^d, Eric Rozet^e, Vincent Delatour^f, Marie-Christine Carlier^g, Anne-Marie Hanser^h, Etienne Cavalierⁱ, Marc Froissart^j, and Jean-Paul Cristol^d
On behalf of the Société Française de Biologie Clinique¹

^a Biochimie Métabolique, Groupe Hospitalier Pitié-Salpêtrière, APHP, Paris, France

^b Nephrology–Dialysis–Transplantation, University of Liège, CHU Sart Tilman, Liège, Belgium

^c Biochimie, CHU Bichat, APHP, Paris, France

^d Biochimie, CHU Lapeyronie, Montpellier, France

^e Analytical Chemistry Laboratory, CIRM, University of Liège, Liège, Belgium

^f Laboratoire National de Métrologie et d'Essais, Paris, France

^g Biochimie, Hôpitaux de Lyon Sud, Lyon, France

^h Biochimie, Hospices civils, Colmar, France

ⁱ Clinical Chemistry, University of Liège, CHU Sart Tilman, Liège, Belgium

^j Physiologie Rénale, Hôpital Européen Georges Pompidou, APHP, Paris, France

ARTICLE INFO

Article history:

Received 5 July 2011

Received in revised form 12 July 2011

Accepted 13 July 2011

Available online xxxx

Keywords:

Creatinine

IDMS traceable

Standardization

ABSTRACT

Chronic kidney disease definition is based on glomerular filtration rate (GFR) estimations which are derived from creatinine-based equations. The accuracy of GFR estimation is thus largely dependent of those of serum creatinine assays. International recommendations highlight the need for traceable creatinine assays. The French Society of Clinical Biochemistry conducted a study for measuring accuracy of creatinine enzymatic methods. This evaluation involved 25 clinical laboratories. Creatinine was measured in serum pools ranging from $35.9 \pm 0.9 \mu\text{mol/L}$ to $174.5 \pm 3.1 \mu\text{mol/L}$ (IDMS determination) using 12 creatinine enzymatic methods. For all creatinine values greater than $74.4 \pm 1.4 \mu\text{mol/L}$, the bias and imprecision did not exceed 5% and 5.9%, respectively. For the lowest value ($35.9 \pm 0.9 \mu\text{mol/L}$), the bias ranged from -1.8 to 9.9% (with one exception). At this level, the imprecision ranged from 1.9 to 7.8%. The true performances of the assays (couples of bias and relative standard deviation), were evaluated using Monte-Carlo simulations. Most of the assays fall within the maximum Total Error of 12% at all concentrations. This study demonstrates substantial improvements in the calibration, traceability and precision of the enzymatic methods, reaching the NKDEP recommendations. Moreover, most of these assays allowed accurate creatinine measurements for creatinine levels lower than $40 \mu\text{mol/L}$.

© 2011 Elsevier B.V. All rights reserved.

Abbreviations: CAP, College of American Pathologists; CKD, chronic kidney disease; EC4, European Communities of Clinical Chemistry; GC-IDMS, isotope dilution-gas chromatography mass spectrometry; GFR, glomerular filtration rate; IDMS, isotope dilution mass spectrometry; KDIGO, Kidney Disease Improving Global Outcomes; K/DOQI, Kidney Disease Outcome Quality Initiative; LAM, liquid assayed multiquant; LNE, Laboratoire National de Métrologie et d'Essais; MDRD, Modification of the Diet in Renal Disease; NIST, National Institute of Standards and Technology; NKDEP, National Kidney Disease Education Program; NKF, National Kidney Foundation; RSD, relative standard deviation; sCr, serum creatinine; SFBC, Société Française de Biologie Clinique; SRM, standard reference material.

* Corresponding author at: Service de Dialyse, CHU Sart Tilman, 4000 Liège, Belgium. Tel.: +32 4 3667317; fax: +32 4 3667205.

E-mail address: pierre_delanaye@yahoo.fr (P. Delanaye).

¹ Affiliations of collaborators (Members of the SFBC group “Biologie des Fonctions rénales”): Zakia Ait Djafer, CHU de Brabois, Nancy; Yann Barguil, CH, Nouméa; Guy Bechis, Hôpital Nord, Marseille; Daniel Bouige, CH, Le Havre; Anne Boutten, CHU Bichat, Paris; Michèle Cambillau, Hôpital Européen Georges Pompidou, Paris; Marie-Christine Carlier, CH Lyon Sud, Lyon; Annette Chamson, Hôpital Nord, Saint Etienne; Jean-Paul Cristol, Hôpital Lapeyronie, Montpellier; Jacques Delavenne, CH, Compiègne; Gérard Desch, CH, Avignon; Soraya Fellahi, Hôpital Tenon, Paris; Sabine Fradin, CHU Côte de Nacre, Caen; Marc Froissart, Hôpital Européen Georges Pompidou, Paris; Marie-Madeleine Galteau, Centre de Médecine Préventive, Vandoeuvre les Nancy; Fabrice Guerber, Laboratoire ELIBIO, La Mûre, Anne-Marie Hanser, Hospices Civils, Colmar; Jean-Philippe Haymann, Hôpital Tenon, Paris; Vincent Larroque, CH A Gayraud, Carcassonne; Marie-Noëlle Letouzey, CH, Villefranche sur Saône; Marie Monge, Pasteur Cerba, Cergy Pontoise; Jean-Gabriel Paul, CHG, Boulogne sur Mer; Laurence Piéroni, Groupe Hospitalier Pitié-Salpêtrière, Paris; Anne Pottier, Laboratoire ELIBIO, La Mûre; Christophe Renard, Hôpital d'Instruction des Armées, Clamart; Sophie Séronie-Vivien, Institut Claudius Regaud, Toulouse.

0009-8981/\$ – see front matter © 2011 Elsevier B.V. All rights reserved.

doi:[10.1016/j.cca.2011.07.012](https://doi.org/10.1016/j.cca.2011.07.012)

Please cite this article as: Piéroni L, et al, A multicentric evaluation of IDMS-traceable creatinine enzymatic assays, Clin Chim Acta (2011), doi:[10.1016/j.cca.2011.07.012](https://doi.org/10.1016/j.cca.2011.07.012)

1. Introduction

Chronic kidney disease (CKD) is a major public health problem in developed countries. Since the first proposal for CKD definition from the National Kidney Foundation (NKF) Kidney Disease Outcome Quality Initiative (K/DOQI) in 2002 [1], the concept of a severity-based international classification of renal diseases has been largely accepted worldwide. Thus, an update for a global definition was proposed in 2005 by the Kidney Disease Improving Global Outcomes (KDIGO) foundation [2]. CKD staging directly relies on GFR values which are mostly estimated by creatinine-based predictive formulas. As a consequence, the accuracy of GFR estimation, and by the way, those of serum creatinine (sCr) assays, has become a major topic, common to clinicians and laboratories.

However, as soon as the K/DOQI classification was proposed, limitations of sCr measurement and the need for standardization have been highlighted [3]. In 2002, the “Creatinine French Working Group” demonstrated that a very high inter-assay variation persisted and that a standardization of the calibration procedure did not reduce enough the variability of sCr assays to allow reliable use of estimated GFR [4]. By the same way, a study performed in 2003 by the College of American Pathologists (CAP) evaluated on fresh frozen serum with a sCr concentration 79.7 $\mu\text{mol/L}$ (0.902 mg/dL), to verify commutability among sCr measurement methods. They showed that a significant bias versus isotope dilution mass spectrometry (IDMS) value persisted [5]. More recently, the EC4 (European Communities of Clinical Chemistry) creatinine standardization working group showed that inter-laboratory variation persisted in reference material prepared from human fresh serum with sCr at a concentration of 76 $\mu\text{mol/L}$ [6].

Since 2006, the “Laboratory Working Group of the National Kidney Disease Education Program” (NKDEP) have described recommendations for improving GFR estimation with guidelines for measuring serum creatinine [7], highlighting the need for traceable and reproducible methods for sCr measurement. The Laboratory Working Group recommended the recalibration of sCr methods in order to be traceable to IDMS. After recalibration, the desirable total error goal should be less than 7.6%, at a sCr concentration of 88.4 $\mu\text{mol/L}$ (1 mg/dL), to ensure a maximum relative 10% increase in the root mean square error of estimated GFR. They also recommended a serum reference material with a sCr target value of 88.4 $\mu\text{mol/L}$, that corresponds to an average GFR of 60 mL/min/1.73 m², to be developed. To meet the latter, the National Institute of Standards and Technology (NIST) has developed a commutable standard reference material (SRM) 967, with sCr at two concentrations, 66.5 $\mu\text{mol/L}$ (0.75 mg/dL) and 346.2 $\mu\text{mol/L}$ (3.93 mg/dL), measured with both GC-MS and LC-MS, in frozen human serum [8]. This SRM is now available to laboratories and manufacturers for the calibration and evaluation of routine clinical sCr methods.

The calibration of sCr methods improved the accuracy of GFR estimation and Murthy et al. recommended that the GFR would be estimated below 60 mL/min/1.73 m² only, unless the method has been calibrated [9].

Some manufacturers adjusted the calibration to reduce the contribution of non-creatinine chromogens in the alkaline-picrate-based methods, primarily by subtracting a constant value as representing the average non-creatinine-dependent signal (so-called “compensated” methods). However, the use of these assays is not accurate on an individual patient basis because of the variability of these chromogens. The evaluation of GFR with such methods in everyday practice for children, elderly patients, pregnant women, cancer or anorectic patients who have low creatinine concentrations could lead to a huge overestimation of renal function [10,11]. Furthermore, the measurement of creatinine in urinary samples, which contain no non-creatinine chromogens, does not need a compensation of the calibration. Consequently, assays based on enzymatic reactions that are much less susceptible to interfere with

non-creatinine chromogens may provide more reliable estimations of GFR.

Since the publication of recommendations, efforts have been made by manufacturers to improve the performances of their creatinine measurement methods, particularly with developing enzymatic methods. To our knowledge, no study has evaluated the performances of enzymatic methods for creatinine assay, after the publication of NKDEP recommendations and the development of SRM. The purpose of SFBC (Société Française de Biologie Clinique) was to check for accuracy of creatinine enzymatic methods commercially available in France.

2. Material and methods

2.1. Experimental design

This evaluation was performed in 2008 and involved 25 clinical laboratories located in 23 French public hospitals and in 2 French private clinical laboratories. sCr was measured in a panel of 5 frozen serum pools and in a Liquid Assayed Multiqual (LAM) Control Level 1 from BIORAD using 12 creatinine enzymatic methods each corresponding to a combination of a reagent and an analyzer. Each method was evaluated in 3 different laboratories.

Sample pools were sent frozen along with LAM Control level 1 (batch number 45571, given to be IDMS traceable to 69 $\mu\text{mol/L}$) by Pasteur Cerba laboratory (MM) to each laboratory where they were stored at $-20\text{ }^{\circ}\text{C}$ until analysis. All data were centralized for statistical analysis at Hôpital Lapeyronnie, Montpellier, France (ASB) and CHU Sart Tilman, Liège, Belgium (ER).

2.2. Serum samples

A panel of five serum pools was prepared in a central laboratory (Pr Galteau, Centre de Médecine Préventive, Nancy, France) in order to cover a wide range of creatinine concentrations from 40 to 200 $\mu\text{mol/L}$, as previously described [4]. During the preparation, sCr was measured using the local non compensated Jaffe assay (Olympus AU 640 using Olympus reagents).

2.3. Principle of enzymatic creatinine methods and details of reagent-analyzer combinations

We evaluated 10 creatinine enzymatic commercially available reagents on 9 analyzers. The main characteristics of the 10 reagents and of the 12 reagent-analyzer combinations are summarized in Table 1. Enzymatic methods used either creatininase (creatinine amidohydrolase EC 3.5.2.10) or creatinine deiminase (creatinine iminohydrolase EC 3.5.4.21).

The first type of enzymatic methods uses a reaction scheme to convert creatinine in sarcosine with the aid of creatininase and creatinase (creatinine amidohydrolase EC 3.5.3.3) and release of hydrogen peroxide with the aid of sarcosine oxidase. In the presence of peroxidase, the liberated hydrogen peroxide reacts by quantitative oxidation with a leuco dye and is measured via an appropriate colorimetric reaction.

In the second type of enzymatic methods, creatinine is converted by creatinine deiminase to ammonia and N-methylhydantoin. The ammonia is subsequently assayed by use of alpha-oxoglutarate and glutamate dehydrogenase. The reduction in absorbance at 340 nm, caused by oxidation of NADPH, is proportional to the ammonia concentration released by creatinine.

2.4. Pre-analytical step

Both pool samples and LAM Control level 1 were thawed at room temperature 1 h before analysis, homogenized by gentle inversion

Table 1
Main characteristics of the 12 automated sCr method–analyzer combinations.

Analyzers	Reagents	Enzymes	Chromogen	Reading mode	λ (nm)	Calibrators	Reference materials for standardization
Beckman Coulter LX20, DxCG600	Sentinel Diagnostics	Creatininase	4-aminoantipyrine + N-ethyl-N-sulfo-propyl-M-toluidine	2 points	546	Human serum 0–333.5 $\mu\text{mol/L}$	NIST SRM 967
Olympus AU2700	Olympus	Creatininase	N-(3-sulfo-propyl)-3-methoxy-5-methylaniline	End point	600/700	Human serum 0–220 $\mu\text{mol/L}$	NIST SRM 967
Siemens RXL	Siemens	Creatininase	4-aminophenazone + 2,4,6-triiodo-3-hydroxybenzoic acid	End point	540/700	Bovine serum 0–983–1949 $\mu\text{mol/L}$	NIST SRM 914 verified by NIST SRM 967
Abbott Architect Ci8200	Abbott	Creatininase	4-aminoantipyrine + N-ethyl-N-sulfo-propyl-m-toluidine	End point	548/700	Human serum 0–326 $\mu\text{mol/L}$	NIST SRM 967
Roche Diagnostics Modular	Roche Creatinine plus	Creatininase	4-aminophenazone + 2,4,6-triiodo-3-hydroxybenzoic acid	End point	546/700	Human serum 0–380 $\mu\text{mol/L}$	GC-IDMS
Roche Diagnostics Modular	Diasys	Creatininase	4-aminoantipyrine + 2,4,6-triiodo-3-hydroxybenzoic acid	2 points	546	Human serum 0–380 $\mu\text{mol/L}$	GC-IDMS
Roche Diagnostics Cobas 6000	Roche	Creatininase	4-aminoantipyrine + 2,4,6-triiodo-3-hydroxybenzoic acid	End point	546/700	Human serum 0–300 $\mu\text{mol/L}$	SRM 909b IDMS
Olympus AU 2700	Randox	Creatinine deiminase	N-methylhydantoine + ammoniac	2 points	340/410	Human serum 0–128 $\mu\text{mol/L}$	GC-IDMS
Olympus AU 2700	Diasys Creatinine PAP FS*	Creatininase	4-aminoantipyrine + 2,4,6-triiodo-3-hydroxybenzoic acid	End point	540/700	Human serum 0–380 $\mu\text{mol/L}$	GC-IDMS
Thermo Fisher KoneLab	Kone	Creatininase	4-aminophenazone + 2,4,6-triiodo-3-hydroxybenzoic acid	End point	550	Bovine serum 0–252 $\mu\text{mol/L}$	NIST SRM 967
Siemens Advia 1800	Siemens	Creatinine deiminase	N-methylhydantoine + ammoniac	Cinetics	340/410	Bovine serum 0–862 $\mu\text{mol/L}$	NIST SRM 914 HPLC
Ortho-Clinical Diagnostics Fusion 5.1 FS	OrthoClinical Diagnostics	Creatininase	2-(3,5-dimethoxy-4-hydroxy phenyl)-4,5-bis(4-dimethylamino phenyl) imidazole	2 points	670	Bovine serum 44–133–1167 $\mu\text{mol/L}$	NIST SRM 914

and centrifuged. After thawing, LAM Control level 1 flasks were stored at 4 °C for the next days of analysis.

2.5. Preliminary period

Imprecision studies were conducted using LAM Control level 1 (batch number 45571, IDMS-traceable). Within-run (30 replicates) and between-run (3 separate runs per day for 10 consecutive days) values were determined in each laboratory ($n = 30$).

Methods not available in the lab's routine practice were implemented by the manufacturer and submitted to a familiarization period including intra-assay reproducibility ($n = 20$ determinations) on the lowest control level provided by the manufacturer.

2.6. Study period

Determinations of the panel were performed on a single batch of reagents after a new calibration and validation of the performance of the method using local internal quality control procedures. The 5 pools were assayed on 3 separate runs per day; LAM Control level 1 was run at the start and the end of the run as samples to verify the run. This test was conducted over 3 consecutive days (1 aliquot thawed each day).

2.7. Reference method

Isotope dilution-gas chromatography mass spectrometry (GC-IDMS) was used as the reference method. Determinations were performed by the Laboratoire National de Métrologie (VD, Paris, France) as previously described [12].

2.8. Statistical analysis

For each reagent–analyzer combinations and pool, evaluation of aberrant data (outliers) was performed using mean reagent–analyzer combination result plus/minus 2 standard deviations. Results were

discarded only if in addition to the statistical evaluation, analytical reasons were determined.

For each laboratory, the average value for each level was calculated. Then these laboratory mean values were grouped per reagent–analyzer combination. Since the trueness was targeted using reference method, plots against the target value could be constructed for each pool level (Fig. 1). Analytical performance was assessed at three levels: bias from the reference value, between-laboratory and within laboratory variation.

The following hierarchical mixed linear model was fitted on all the data to evaluate the fixed effects of creatinine concentrations of the pools and the random effects of the assays, laboratories and runs as well as the interaction of the concentration levels with the assays:

$$y_{ijklm} = \mu + \alpha_i + \beta_j + \gamma_{kij} + \eta_{l[j,k]} + \alpha_i\beta_j + \varepsilon_{ijklm}.$$

μ is the overall mean. α_i represents the fixed effect of the concentration levels or creatinine pools ($i: 1$ to 5), β_j represents the random effect of the assays ($j: 1$ to 16), γ_{kij} is the random effect of the laboratories nested into the assays ($k: 1$ to 3), $\eta_{l[j,k]}$ is the random effect of the runs nested into the assays and laboratories ($l: 1$ to 3) and $\alpha_i\beta_j$ is the concentration–assay interaction. All the random effects were assumed coming from independent central normal distributions. ε_{ijklm} is the residual error (or repeatability) assumed normally distributed $N(0, \sigma_\varepsilon^2)$. For the fixed effect, significance level was settled at 5%.

On each pool of creatinine ($i = 1$ to 5), the following random ANOVA III hierarchical model was fitted in order to evaluate the importance in terms of % Relative Standard Deviation (RSD) of each factor assessed, namely: assay kit, laboratory, run and repeatability (or residual error):

$$y_{ijklm} = \mu + \beta_j + \gamma_{kij} + \eta_{l[j,k]} + \varepsilon_{ijklm}.$$

Then, computations of relative biases (%) and RSDs (%) for all assays at all pools of creatinine were computed using the results obtained by the reference laboratory as “true” creatinine concentrations

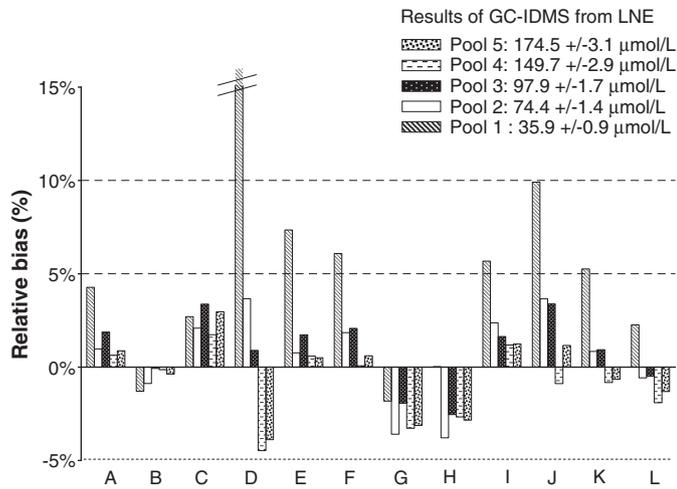


Fig. 1. Relative bias according to GC-IDMS creatinine in 5 pools ranging from 36 to 175 $\mu\text{mol/L}$. Twelve methods corresponding to a reagent/analyzer combination were evaluated. A: Roche Modular/Roche; B: Roche Cobas/Roche; C: Olympus/Randox; D: OrthoClinical/OrthoClinical; E: Olympus/Diasys; F: Roche Modular/Diasys; G: Siemens RXL/Siemens; H: Abbott/Abbott; I: Beckman Coulter/Sentinel Diagnostics; J: ThermoFischerKoneLab/Kone; K: Olympus/Olympus; L: Siemens Advia/Siemens.

for the bias estimations. JMP v7.0.1 (SAS Institute, Cary, NC) software was used to perform these computations.

Based on these results all the possible true couples of bias and RSD values for each method at all pools giving 95 times out of 100 the

estimated values obtained in the study were searched by Monte-Carlo simulations. Simultaneously, the true region of couples (bias, RSD) providing at most 8% Total Error and 12% Total Error was computed and all these elements were graphed for interpretations. These graphs plotted systematic bias and imprecision of the reagent-analyzer combination into a parabolic curve initially developed by Myers et al. (Fig. 2). These analysis and graphs were performed with R v2.9.1 (CRAN, <http://cran.r-project.org>).

3. Results

3.1. Target values for the five pools investigated

According to GC-IDMS, the true creatinine levels of the five pools investigated were $35.9 \pm 0.9 \mu\text{mol/L}$, $74.4 \pm 1.4 \mu\text{mol/L}$, $97.9 \pm 1.7 \mu\text{mol/L}$, $149.7 \pm 2.9 \mu\text{mol/L}$ and $174.5 \pm 3.1 \mu\text{mol/L}$ for pools 1, 2, 3, 4 and 5, respectively.

3.2. Sources of variability

The hierarchical linear mixed model fitted well the whole set of data as shown by the high value of the adjusted R^2 equal to 0.998. The concentration levels of creatinine were found highly statistically significant ($p\text{-value} < 0.0001$), revealing that the results of the assays were dependent of the true concentrations of the pools. Furthermore the interaction coefficient between concentration levels and assays was also highly significant ($p\text{-value} < 0.0001$) demonstrating that all the assays were not responding similarly when the creatinine

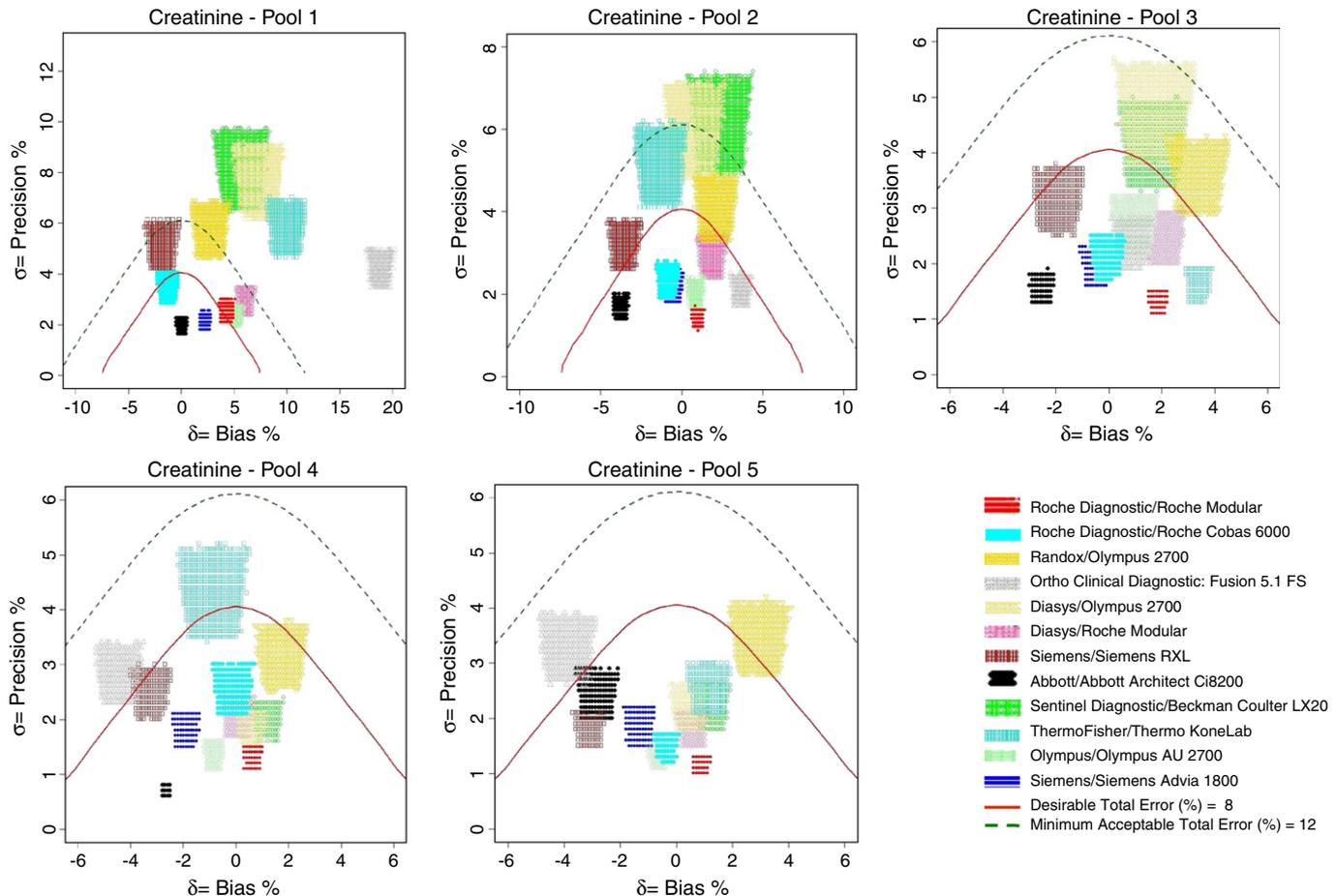


Fig. 2. Monte-Carlo simulations representing 95% probable regions of true performances of the assays for different pools of creatinine. Monte-Carlo simulations represent the true couples of bias and RSD values that can give a 95% probability of the bias and RSD observed in the study. Continuous curve: region defined by true values of bias and RSD with a maximum of 8% Total Error. Dashed curve: region defined by true values of bias and RSD with a maximum of 12% Total Error.

concentration increases. The random effects of assays, lab runs and repeatability measured as RSDs showed that the assays were responsible of the most variability in the results (RSD = 1.5%) together with the laboratories (RSD = 1.5%), whereas the runs and repeatability parts were the smallest with similar levels (RSD of 0.5 and 0.4%, respectively). This illustrates that the results of creatinine can be highly different from one assay to the other, as well as from one laboratory to the other (data not shown).

3.3. Analytical performances (imprecision and bias according to the reference method)

To deepen the analysis, the bias of each assay was computed by comparison to the results obtained by the national reference laboratory as well as the imprecision of each assay. The results of imprecision for each pool of creatinine are illustrated in Table 2.

For the lowest pool ($35.9 \pm 0.9 \mu\text{mol/L}$) the bias ranged from -1.3 to 18.8% and the imprecision from 1.9 to 7.3% . The minimum and maximum biases were observed for the assays Abbott on Architect and OrthoClinicalDiagnostics on Fusion 5.1, respectively. The maximum imprecision was observed for the Sentinel reagents on Beckman Coulter LX. For all other pools, the observed bias never exceeded 5% , ranging from -4.5% (OrthoClinicalDiagnostics, pool 4) to 3.8% (OrthoClinicalDiagnostics, pool 2). Similarly, the imprecision ranged from 0.70% (Abbott Architect, pool 4) to 5.9% (Beckman Coulter LX, pool 2).

3.4. True performances of the assays defined by bias and RSD

Based on these observed performances of the assays expressed in terms of couples of bias and RSD, the true performance of the assays were searched by Monte-Carlo simulations. They represent all the true couples of bias and RSD values of the assays that can give the observed values of bias and RSD of the study 95 times out of 100. The results are shown in Fig. 2 for each pool of creatinine. Furthermore, the regions of methods defined by their true values of bias and RSD that have a maximum of 8% Total Error (desirable creatinine assays: continuous line) and a maximum of 12% Total Error (acceptable creatinine assays: dotted line) as proposed by Myers are depicted on these graphs [7]. The assays that answer these specifications of maximum 8% and 12% Total Error are thus those falling inside these regions. As shown in Fig. 2, far from all the assays fall within the maximum Total Error of 12% at all the concentrations of creatinine. Only four reagent/analyzer combinations do not fit this requirement at the first level of $35 \mu\text{mol/L}$ (Ortho Clinical/Ortho Clinical Fusion 5.1, ThermoFisher/ThermoFisher Konelab, Sentinel Diagnostics/Beckman Coulter LX20 and Diasys/Olympus AU 2700).

At the second level of $74 \mu\text{mol/L}$, three of these four reagent/analyzer combinations are on the 12% borderline but all the other meet the specifications.

4. Discussion

In this study, we have shown and illustrated the substantial improvements in the calibration, traceability and precision for most of the enzymatic methods used for creatinine measurement. This study demonstrates for the first time that NKDEP recommendations could be achieved using enzymatic creatinine assays. Moreover, most of these assays allowed accurate creatinine measurements for creatinine levels lower than $40 \mu\text{mol/L}$.

The calibration matter for creatinine is especially relevant because creatinine is more and more automatically used by laboratories to estimate GFR, notably with the Modification of the Diet in Renal Disease (MDRD) study equations [7,13–16]. The impact of the calibration on the MDRD results is far from negligible, especially in the high GFR range [7,9,17–19]. The MDRD study equation has even been modified to be useful with IDMS traceable creatinine [13,15]. The need for a traceable, reproducible, method for sCr measurement could be further enhanced using CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration) formula taking into account low creatinine value (lower than $62 \mu\text{mol/L}$ for women) [20].

Several authors have illustrated the lack of standardization for creatinine measurement until now. One of the most descriptive studies is certainly the study published in 2005 by Miller on behalf of College of American Pathologists [5]. Comparing fifty methods of creatinine measurement in 5624 American laboratories with a reference method (IDMS), these authors find large discrepancies and bias between the methods. Creatinine results with different enzymatic methods differ from the reference method from -1.8 to $17.7 \mu\text{mol/L}$ (-0.02 to 0.2 mg/dL). In this study, only one level of creatinine was studied and the statistical weight of methods differ from each other allowing the number of laboratories using a method-analyzer combination (for example, the picrate method from Roche is used by 78 laboratories whereas the picrate method from Olympus is only used in 12 laboratories) [5]. Other studies in France [4] and United Kingdom [18] have also underlined the lack of standardization and the high inter-assay and inter-laboratory CV.

Another important study on this topic has been published in 2006 by Myers on behalf of the NKDEP Laboratory Working Group [7]. This working group has determined the total analytical error in creatinine that may be considered as acceptable (total error budget). This total error combines the systematic bias, due to differences of calibration, and the random measurement error which includes within-laboratory

Table 2

Imprecision between laboratories. Imprecision expressed as relative SD% according to different creatinine pools and reagent/analyzer combinations.

		Roche	Roche	Randox	OrthoClinical	Diasys	Diasys	Siemens	Abbott	Sentinel	ThermoFisher	Olympus	Siemens
		Modular	Cobas	Olympus	Fusion 5,1FS	Olympus	Roche	RXL	Architect	Beckman	KoneLab	AU 2700	Advia
			6000	AU		AU	Modular			LX			1800
Pool 1	Mean	37.43	35.43	36.87	42.66	38.53	38.08	35.24	35.90	37.93	39.45	37.78	36.71
	SD	0.91	1.17	2.04	1.71	2.82	1.08	1.73	0.66	2.95	2.19	0.84	0.75
	CV%	2.43	3.31	5.53	4.01	7.33	2.84	4.93	1.85	7.78	5.57	2.23	2.06
Pool 2	Mean	75.12	73.74	75.95	77.12	74.96	75.77	71.72	71.58	76.16	77.12	75.02	73.97
	SD	1.01	1.65	2.98	1.55	4.28	2.14	2.19	1.16	4.47	1.75	1.46	1.54
	CV%	1.34	2.24	3.93	2.01	5.71	2.82	3.06	1.62	5.87	2.28	1.94	2.08
Pool 3	Mean	99.73	97.83	101.20	98.77	99.59	99.92	96.00	95.40	99.50	101.20	98.81	97.41
	SD	1.23	1.99	3.47	2.26	4.50	2.38	2.89	1.45	3.97	1.59	2.59	1.85
	CV%	1.24	2.04	3.43	2.28	4.52	2.38	3.01	1.52	3.99	1.57	2.62	1.89
Pool 4	Mean	150.64	149.49	152.3	142.99	150.58	149.80	144.79	145.70	151.47	148.37	148.46	146.83
	SD	1.86	3.65	4.59	3.89	2.63	2.96	3.48	0.98	2.89	6.15	1.95	2.55
	CV%	1.23	2.44	3.01	2.72	1.74	1.98	2.40	0.67	1.91	4.14	1.31	1.74
Pool 5	Mean	176.02	173.84	179.67	167.71	175.37	175.55	169.04	169.54	176.66	176.51	173.36	172.21
	SD	1.94	2.48	5.98	5.27	3.71	3.08	2.86	4.01	3.73	4.29	2.12	3.08
	CV%	1.10	1.43	3.33	3.14	2.11	1.75	1.69	2.36	2.11	2.43	1.22	1.78

effect, between-laboratory random variability in day-to-day calibration and specimen-specific effect. Myers calculated desirable total error of 7.6%. This assumption is based on the intra- and inter-individual biological variation of creatinine [7]. This concept was reported in 2008 by Delanghe who has conducted a similar study than Miller in Europe. Three levels of creatinine were analyzed by 189 laboratories. Among these laboratories, 30% were measuring creatinine with an enzymatic method but only 3 reagent-analyzer combinations. In this study, the budget of the total error was especially consumed by the systematic bias. This fact is especially relevant in enzymatic methods because the random measurement error is low with these methods [6]. Indeed, analytical performances of the enzymatic creatinine measurement are clearly better than for the classical Jaffe methods [10,11]. Interferences such as bilirubin or hemolysis also seem less important with enzymatic methods, as it has been elegantly shown by Cobbaert [10].

According to these previous data, we have focused our work on these enzymatic methods. Moreover, it is evident that the impact of the analytical error in creatinine measurement is only relevant in the relatively low levels of creatinine. We have thus studied five creatinine levels insisting on the low levels of creatinine. This study has been conducted in 2008. At this time, the creatinine calibration problems were well known, had been extensively illustrated in the literature and were well known by all scientists, clinical biologists, nephrologists and manufacturers [11,21]. Most of the manufacturers have claimed to have improved their calibration to the IDMS [21,22]. The main finding of our study is that enzymatic methods have improved their calibration to IDMS creatinine and reached a relative bias lower than 5% at level up to 70 $\mu\text{mol/L}$ as recommended by NKDEP. Furthermore, using a low creatinine level (36 $\mu\text{mol/L}$), the relative bias is lower than 10% except for dry chemistry. The marked bias differences between OrthoClinical Diagnostics reagents and other creatinase assays strongly suggest that the positive bias in the dry chemistry method could be due to difference in calibration. This observation confirms and extends Delanghe's previous report showing a constant overestimation using dry chemistry. Interestingly, after Delanghe's work, performed in 2005 [6], OrthoClinical Diagnostics has recalibrated its assay leading to an IDMS traceable creatinine. This process prevents the overestimation in value greater than 77 $\mu\text{mol/L}$ (0.87 mg/dL), but has to be further improved for low values. It could be hypothesized that this specific dry chemistry bias could be, at least in part, due to a matrix effect or use of bovine materials.

Our study clearly evidences an improvement in creatinine measurements in comparison to other studies in the recent past. Indeed, our results demonstrated that the majority of enzymatic methods reach the total analytical error of 8% as demonstrated by the Monte-Carlo simulation. It seems fundamental to underline that such results are obtained for creatinine values as low as 36 $\mu\text{mol/L}$ (0.41 mg/dL).

Our study has several strengths. We have studied more than 12 reagent-analyzer combinations (25 laboratories have participated) on five different levels of creatinine. Our comparison has been realized against a valid and strong reference method and we used frozen human sera to avoid matrix effect [23,24]. Contrary to other studies, each method has been studied and validated in the same number of laboratories and with the same methodology which makes easier statistical interpretations [5,6]. However, there are some limitations. We have tried to be exhaustive in the choice of the enzymatic methods but some methods have been omitted either because not enough laboratories using this method could be included or because the methods were still not available when the study has been designed. We have validated creatinine measurement in the low concentration range. However, we have no proof that precision and calibration are linear and performing well between 0 and 36 $\mu\text{mol/L}$. These low values are however only important in pediatrics [10]. As previously shown, we confirm that precision (random error measurement) of enzymatic methods is high for most of the methods. However, for the first time, we

can affirm that several enzymatic methods are adequately calibrated against IDMS, substantially improving standardization and traceability of this renal marker. This "good news" will have important consequences on the precision of the estimated GFR by the creatinine-based equations [16].

Acknowledgments

A research grant from the Belgium National Fund for Scientific Research (F.R.S-FNRS) to E. Rozet is gratefully acknowledged.

References

- [1] National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am J Kidney Dis* 2002;39: S1–S266.
- [2] Levey AS, Eckardt KU, Tsukamoto Y, et al. Definition and classification of chronic kidney disease: a position statement from Kidney Disease: Improving Global Outcomes (KDIGO). *Kidney Int* 2005;67:2089–100.
- [3] Coresh J, Eknoyan G, Levey AS. Estimating the prevalence of low glomerular filtration rate requires attention to the creatinine assay calibration. *J Am Soc Nephrol* 2002;13:2811–2.
- [4] Seronie-Vivien S, Galteau MM, Carlier MC, et al. Impact of standardized calibration on the inter-assay variation of 14 automated assays for the measurement of creatinine in human serum. *Clin Chem Lab Med* 2005;43:1227–33.
- [5] Miller WG, Myers GL, Ashwood ER, et al. Creatinine measurement: state of the art in accuracy and interlaboratory harmonization. *Arch Pathol Lab Med* 2005;129: 297–304.
- [6] Delanghe JR, Cobbaert C, Galteau MM, et al. Trueness verification of actual creatinine assays in the European market demonstrates a disappointing variability that needs substantial improvement. An international study in the framework of the EC4 creatinine standardization working group. *Clin Chem Lab Med* 2008;46:1319–25.
- [7] Myers GL, Miller WG, Coresh J, et al. Recommendations for improving serum creatinine measurement: a report from the laboratory working group of the national kidney disease education program. *Clin Chem* 2006;52:5–18.
- [8] Dodder NG, Tai SS, Sniegoski LT, Zhang NF, Welch MJ. Certification of creatinine in a human serum reference material by GC–MS and LC–MS. *Clin Chem* 2007;53:1694–9.
- [9] Murthy K, Stevens LA, Stark PC, Levey AS. Variation in the serum creatinine assay calibration: a practical application to glomerular filtration rate estimation. *Kidney Int* 2005;68:1884–7.
- [10] Cobbaert CM, Baadenhuijsen H, Weykamp CW. Prime time for enzymatic creatinine methods in pediatrics. *Clin Chem* 2009;55:549–58.
- [11] Panteghini M. Enzymatic assays for creatinine: time for action. *Scand J Clin Lab Invest Suppl* 2008;241:84–8.
- [12] Thienpont LM, Van Landuyt KG, Stockl D, De Leenheer AP. Candidate reference method for determining serum creatinine by isocratic HPLC: validation with isotope dilution gas chromatography–mass spectrometry and application for accuracy assessment of routine test kits. *Clin Chem* 1995;41:995–1003.
- [13] Levey AS, Coresh J, Greene T, 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.
- [14] Levey AS, Stevens LA, Hostetter T. Automatic reporting of estimated glomerular filtration rate—just what the doctor ordered. *Clin Chem* 2006;52:2188–93.
- [15] Levey AS, Coresh J, Greene T, 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.
- [16] Stevens LA, Manzi J, Levey AS, 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.
- [17] Delanaye P, Cavalier E, Chapelle JP, Krzesinski JM. Importance of the creatinine calibration in the estimation of GFR by MDRD equation. *Nephrol Dial Transplant* 2006;21:1130.
- [18] Lawson N, Lang T, Broughton A, Prinsloo P, Turner C, Marenah C. Creatinine assays: time for action? *Ann Clin Biochem* 2002;39:599–602.
- [19] Van Biesen W, Vanholder R, Veys N, et al. The importance of standardization of creatinine in the implementation of guidelines and recommendations for CKD: implications for CKD management programmes. *Nephrol Dial Transplant* 2005;21:77–83.
- [20] Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009;150:604–12.
- [21] Kytzia HJ. How to implement traceability of creatinine results: a manufacturer's experience. *Scand J Clin Lab Invest Suppl* 2008;241:64–6.
- [22] Myers GL. Standardization of serum creatinine measurement: theory and practice. *Scand J Clin Lab Invest Suppl* 2008;241:57–63.
- [23] Stockl D, Libeer JC, Reinauer H, Thienpont LM, De Leenheer AP. Accuracy-based assessment of proficiency testing results with serum from single donations: possibilities and limitations. *Clin Chem* 1996;42:469–70.
- [24] Carobene A, Ferrero C, Ceriotti F, et al. Creatinine measurement proficiency testing: assignment of matrix-adjusted ID GC–MS target values. *Clin Chem* 1997;43:1342–7.