Clinica Chimica Acta 427 (2014) 100–106

Contents lists available at ScienceDirect

Clinica Chimica Acta

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

Evaluation of the impact of standardization process on the quality of serum creatinine determination in Italian laboratories

Anna Carobene [a,*], Ferruccio Ceriotti [a], Ilenia Infusino [b], Erika Frusciante [b], Mauro Panteghini [b]

[a] Diagnostica e Ricerca San Raffaele, Istituto Scientifico Ospedale San Raffaele, Milano, Italy
[b] Centro Interdipartimentale per la Riferibilità Metrologica in Medicina di Laboratorio (CIRME), Università degli Studi, Milano, Italy

ARTICLE INFO

Article history:
Received 25 July 2013
Received in revised form 12 September 2013
Accepted 4 October 2013
Available online 19 October 2013

Keywords:
Standardization
Creatinine
EQAS

ABSTRACT

Background: Creatinine determination in serum is a key indicator of kidney glomerular function. A reference measurement system for its standardization is available and virtually all IVD manufacturers have aligned their assays to this system. In this study, we verified the impact of these standardization efforts on the results of an Italian EQAS involving about 430 laboratories.
Method: We considered data obtained during 2006, 2010 and 2011 schemes of EQAS Prolarit for control materials with target values assigned by a traceable method (enzymatic assay calibrated against the NIST SRM 967).
Results: The results showed a good alignment at concentrations ~170 μmol/L, with 2011 results — except for one method group — well inside the desirable bias (±4%). At higher concentrations, whereas the bias was small in 2010, for some groups using alkaline-picrate (AP) methods it became significantly negative in 2011. The performance seems to worsen when measuring physiologic concentrations, where a significant positive bias (up to ~20%) is shown by most of the AP-based analytical systems. With few exceptions, no evident improvement in individual assay bias was noted from pre- (2006) to post-standardization (2011) periods. The enzymatic method groups were the only always presenting an acceptable bias at all creatinine concentrations, also showing the lowest between-laboratory variability.
Conclusion: Our data seem to indicate that the standardization efforts are still having effects lower than expected. Even taking into consideration that some of the bias may derive from non commutability problems, most of the current "standardized" AP-based methods, at the lower creatinine concentrations, seem to present accuracy problems. This inaccuracy can adversely impact the estimation of GFR by equations and the evaluation of kidney function in pediatrics.

1. Introduction

The measurement of creatinine in serum is one of the most frequently requested tests in clinical laboratories. Over the last decade, its importance in the evaluation of kidney glomerular function has been further increased by recommendations issued by nephrology societies for using equations to estimate the glomerular filtration rate (GFR) [1]. Particularly, the equation developed by the Modification of Diet in Renal Disease (MDRD) study or, if GFR is >60 mL/min, developed by the Chronic Kidney Disease-Epidemiology Collaboration (CKD-EPI) is recommended [2–4]. Recommendations specifically include the requirement for standardization of creatinine measurements before the implementation of GFR estimation by equation [5]. This has resulted in an international consensus for the definition of a reference system for the measurement of creatinine based on the use of higher-order reference materials and reference methods to which the calibration of commercial assays has to be traceable [6]. Accordingly, from 2008 in vitro diagnostic (IVD) manufacturers have (re)calibrated their creatinine measurement systems to the internationally approved reference system and, consequently, clinical laboratories are now exclusively using these standardized methods. This is also in compliance with the European Union (EU) Directive 98/79/EC on IVD medical devices, which among essential requirements for the CE ("Communautés Européennes") mark states that "the traceability of values assigned to calibrators and/or control materials must be assured through available reference measurement procedures and/or available reference materials of a higher order" [7,8].

In 2006, experimental data were obtained both at European and Italian level describing the situation of creatinine measurement in pre-standardization times [9,10]. In a study involving ~170 European laboratories, Delanghe et al. [9] demonstrated that analytical systems from four of the most important IVD manufacturers did not meet traceability targets, especially when physiologic concentrations of serum

Abbreviations: AP, alkaline-picrate; GFR, glomerular filtration rate; MDRD, Modification of Diet in Renal Disease; CKD-EPI, Chronic Kidney Disease-Epidemiology Collaboration; IVD, in vitro diagnostic; CE, Communautés Européennes; EQAS, external quality assessment schemes; 95PI, 95% prediction interval; CIRME, Centre for Metrological Traceability in Laboratory Medicine; NIST, National Institute for Standards and Technology; IDMS, isotopic dilution-mass spectrometry.

* Corresponding author at: Diagnostica e Ricerca San Raffaele, Via Olgettina 60, 20132 Milano, Italy. Tel.: +39 02 26432850; fax: +39 02 26434187.
E-mail address: carobene.anna@hsr.it (A. Carobene).

0009-8981/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cca.2013.10.001

creatinine were measured. A similar situation was confirmed at a national level in a study in which, for physiological concentrations of serum creatinine, the total average overestimation was relevant (>10%), with >20% peaks [10]. Only some analytical systems using methods based on enzymatic principles produced sufficiently accurate results. At higher creatinine concentrations the situation tended to improve, but only at very high concentration levels (~300 μmol/L) a consistent percentage of analytical systems, and therefore of laboratories, succeeded in obtaining results within the acceptable limits of bias [10].

The aim of our work was to check if and how much these standardization efforts have influenced the results provided by laboratories participating in the national EQAS Prolarit. EQAS results of year 2006 (pre-) were compared with those of the biennium 2010–2011 (post-standardization period). Unfortunately the EQAS Prolarit cannot be classified as category 1 or 2 scheme [11] because the commutability of the materials is not fully demonstrated. However a target value was assigned to the EQAS materials for the considered analyte through a reference like procedure performed by an accredited laboratory. Finally, an objective criterion was used to assess the quality of measurement made by laboratories to check their suitability for clinical application. Moreover, being that the characteristics of the control materials are similar with time and the population of the clinical laboratories is stable, the collected data may be able to show the existence of a possible positive trend.

2. Materials and methods

2.1. EQAS and results of participating laboratories

Data utilized in this study were obtained from the results provided by laboratories participating in the EQAS Prolarit, organized by "Diagnostica e Ricerca San Raffaele", Milan. About 430 laboratories from almost all regions of the country participate in the clinical chemistry scheme. The scheme uses 12 lyophilized materials, sent in a single shipment, to be stored at 2–8 °C and measured at bimonthly intervals. EQAS materials are produced by BioRad (Segrate, Italy); declared homogeneity (variability among vials) and stability (degradation in the life's time of the material) are <1%.

For the scope of the study, results of 12 different control materials with creatinine concentrations of about 80, 115, 170 and 210 μmol/L and total protein concentration of about 47, 54, 63 and 72 g/L respectively, distributed in three different cycles (2006, 2010 and 2011) were considered. For every year, four sets of ~400 results were obtained, i.e., 1600 results per year, for a total of more than 4000 data. From these data, only those obtained from homogeneous groups (i.e., platform, reagents and calibrators from the same manufacturer) were retrieved, corresponding to ~250 results per material per year, for a total of ~3000 results. The data derivation was performed by grouping homogeneous groups and analytical principles: Table 1 shows the list of analytical systems included in various method groups. The number and types of analyzers varied slightly, but not substantially, from year to year. The only heterogeneous group, both for instrumentations and reagents, was that defined as "Enzymatic Others". This group consisted of methods based on enzymatic principles, but using reagents and platforms from different manufacturers, with the only exception of Roche Diagnostics that reached a number of participants large enough to allow the definition of a specific group. For each group and for each EQAS material, the mean and inter-laboratory CV, and the bias of the group mean from the target were calculated after outlier elimination (i.e., results deviating from the mean > ± 3SD).

Table 1
Composition of evaluated groups of analytical systems.

Method/ manufacturer group	Analytical systems
Enzymatic Roche	Modular, Cobas, Integra 400 and 800
Enzymatic Others[a]	Abbott Ci8200, Beckman AU400, AU600, AU640, Biotecnica BT3000 and BT3500, Instrumentation Laboratory ILab 600 and 650, Siemens Vista, Advia 1650
AP Beckman	DXC 600 and 800, Synchron CX3, CX7 and CX9
AP Biotecnica	ARCO PC, Targa 2000, 3000 and 10000, BT 2245, BT 3000 Plus, BT 3500
AP Siemens Dimension	Dimension X-Pand, Dimension AR, Dimension RXL
AP Roche Hitachi	Hitachi 704, 717, 911, 912 and 917, Modular and Modular SWA, Cobas 6000 and Cobas c501
AP Instrumentation Laboratory	ILab 300 and 300 plus, ILab 500, 600 and 650
AP Roche Integra	Integra 400, 400 plus and 800
AP Kone	KoneLab 20i, 30i, 60i, 600 and T60
AP Beckman AU	AU400, AU600, AU640, AU680, AU2700 and AU5400
AP ABX	Cobas Mira
AP Siemens Advia	Advia 1650, 1800 and 2400

AP, alkaline-picrate method.
[a] Constituted by systems using enzymatic assays other than Roche.

2.2. Evaluation of commutability of EQAS materials

The commutability of control materials used in this study was verified according to the C53-A document of the Clinical and Laboratory Standards Institute (CLSI) [12]. In particular, the commutability was evaluated by comparing the behavior of materials with that of 30 anonymous leftover samples from single donor [creatinine concentrations: ≤100 μmol/L (n = 10), between 100 and 220 μmol/L (n = 10) and >220 μmol/L (n = 10)] when measured by two analytic methods, using an enzymatic (Cobas c501, Roche Diagnostics) and an alkaline-picrate (AP) analytical principle (Advia 2400, Siemens Healthcare Diagnostics), respectively. The control materials were deemed commutable when their values were inside the 95% prediction interval (95PI) calculated from the Deming regression of the 30 native samples [12].

2.3. Target value assignment to EQAS materials

The target values to selected EQAS materials were assigned by the accredited (ISO 17025 and 15195 standards) reference laboratory of the Centre for Metrological Traceability in Laboratory Medicine (CIRME) of the University of Milan. Measurements were performed using a Roche Cobas c501 platform, the CREP2-Creatinine plus ver. 2 reagent (Roche Diagnostics, cod. 03263991, lot no. 661903) and the two-level matrixed reference material SRM 967a [National Institute for Standards and Technology (NIST)] as calibrator [value assigned by isotopic dilution-mass spectrometry (IDMS) ± combined expanded uncertainty: 74.87 μmol/L ± 1.59 μmol/L and 342.73 μmol/L ± 7.25 μmol/L, respectively] [13]. The analytical performance of the method was previously evaluated [14]. In particular, a correlation with IDMS method gave the following regression equation: CIRME enzymatic method = 0.997 IDMS − 1.47 μmol/L (r = 0.9999).

For target value assignment three different aliquots of each control material were measured in three different runs (9 results in total). The mean of each run and the grand mean (mean of the three means) were calculated. The combined standard uncertainty of each target value was calculated considering both the measurement imprecision and the uncertainty of the value assigned to the calibrator SRM 967, using the formula $\sqrt{CV^2 + uSRM^2}$, where CV is the standard uncertainty derived by the imprecision and the uSRM is the uncertainty of SRM 967 target value (1.06%). From the combined uncertainty, the expanded uncertainty was estimated using a coverage factor of 2 (confidence interval = 95%).

2.4. Criteria for evaluating quality of analytic performance

Criteria based on biologic variability of creatinine in serum [15] were employed to derive both the maximum allowable bias (used to evaluate the performance of the method/manufacturer groups) and the maximum

allowable total error (TE) (used to evaluate the performance of individual laboratories) [16]. Based on these criteria, the maximum allowable bias was ±4.0% (desirable level) and ±6.0% (minimum quality level), and the maximum allowable TE was ±8.9% (desirable level) and ±13.4% (minimum quality level).

2.5. Statistical analysis

All statistics were performed using Microsoft Office Excel 2003, except for Deming regression that was carried out with the SigmaPlot v.12 software (Systat Software, Inc.).

3. Results

When testing commutability, all results related to tested EQAS materials were inside the 95PI of the regression obtained using results from native human serum samples, then demonstrating a similar behavior. As an example, the results of commutability experiment related to the EQAS 2011 materials are shown as supplemental data.

Tables 2 to 5 (top) display the target values (±expanded uncertainty) assigned to 12 EQAS materials used in the study. It has been postulated [17] that the uncertainty of target values assigned to control materials by a reference laboratory should be lower than 20% of the maximum TE allowed in the scheme, otherwise the variability of the target values might influence the evaluation of the performance of participating laboratories. The average uncertainty obtained in our reference laboratory (2.5%) was indeed <20% of the minimum allowable TE goal (±13.4%), thus deemed suitable for the aim of the study.

Tables 2–5 also show results on the selected materials obtained by participating laboratories grouped by method principle/analytical systems during the three evaluated years of EQA. For each group, the mean value, bias (in percentage) of the group mean from the target, inter-laboratory CV and the number of laboratories are shown. For creatinine concentrations of ~170 μmol/L, the results showed an overall good alignment, with all group means for material H 2011 (except AP Beckman group) within the desirable bias goal (±4%) (Table 4). Looking at the materials with higher creatinine concentrations (Table 5), all biases in 2010 were within the desirable limit; however, in 2011 some groups employing AP methods worsened their performance, with the Siemens Advia group showing a relevant underestimation (−8.4%). Overall, EQAS materials with physiologic creatinine concentrations (80–115 μmol/L) gave the worse results, both in terms of interlaboratory CV (very frequently higher than 10%) and bias, with a positive bias (up to 20%) often present for systems using AP methods (Tables 2 and 3).

Unexpectedly, the standardization of most analytical systems did not show a major improvement from the pre- (2006) to the post-standardization period (2010–2011). The only exceptions were represented by AP-based Roche systems (both Hitachi and Integra) and by AP-based Instrumentation Laboratory systems, the latter group, however, still showing in 2011 an apparent relevant systematic error (+12.8%) at creatinine concentrations of 88.4 μmol/L (material B 2011). On the other hand, in the post-standardization period the enzymatic groups always showed an acceptable bias, independently of creatinine concentrations. Moreover, these two groups also showed the lowest CV among laboratories.

Fig. 1a and b summarizes these results showing an overall very limited effect of the standardization effort for the AP-based methods.

As creatinine measurements were performed by individual laboratories participating to the EQAS, we could not exclude that a poor local attitude, in principle not related to the basic performance of the employed analytic system, influenced the quality of obtained results. Consequently, we finally decided to focus on the performance of individual laboratories and the analytic quality of their creatinine measurements by evaluating results of more recently analyzed EQAS materials with physiologic and borderline concentrations [B 2011 (assigned value, 88.4 μmol/L) and D 2011 (assigned value, 123.8 μmol/L)] vs. the desirable goal for TE (±8.9%). Fig. 2 reports as Youden plots the corresponding results by laboratories using enzymatic assays and those using AP assays. Notwithstanding the marked difference in size of two groups, it was evident that the vast majority (87%) of laboratories using systems employing enzymatic assays were able to fulfill the desirable performance, while only one third of laboratories using AP-based systems were able to meet the target.

4. Discussion

4.1. Study limitations

In our study there are two main limitations:

Target value assignment to EQAS materials: The method used to assign the target values (enzymatic assay calibrated against the NIST SRM 967) is not one of the JCTLM listed methods.

Evaluation of commutability of EQAS materials: It was not possible to demonstrate the commutability of our control material in a comparison between the IDMS method and our enzymatic method. For this reason

Table 2
Results for EQAS materials with creatinine concentrations of about 80 μmol/L for the evaluated years.

Group	Material B 2006 — target value: 79.6 μmol/L ± 1.9 μmol/L				Material A 2010 — target value: 71.6 μmol/L ± 1.9 μmol/L				Material B 2011 — target value: 88.4 μmol/L ± 2.1 μmol/L			
	Mean, μmol/L	Bias[a]	CV[b]	n	Mean, μmol/L	Bias	CV	n	Mean, μmol/L	Bias	CV	n
Enzymatic Roche	78.7	−0.9%	1.6%	4	67.2	*−5.7%*	3.9%	8	87.5	−1.2%	5.5%	11
Enzymatic Others	–	–	–	–	69.0	−3.3%	1.6%	8	87.5	−1.2%	5.9%	19
AP Beckman	82.2	3.3%	4.2%	11	68.1	*−4.6%*	14.9%	11	91.9	4.0%	7.7%	5
AP Biotecnica	92.8	**17.3%**	9.9%	12	88.4	**24.0%**	9.8%	9	97.2	**10.0%**	12.2%	14
AP Dimension	84.9	**6.4%**	13.3%	10	77.8	**9.3%**	7.8%	11	94.6	**6.9%**	8.7%	19
AP Roche Hitachi	86.6	**8.7%**	7.1%	26	71.6	0.6%	7.0%	23	92.8	*4.4%*	5.5%	36
AP IL	100.8	**26.4%**	11.5%	19	82.2	**14.9%**	10.3%	27	99.9	**12.8%**	8.7%	46
AP Roche Integra	86.6	**8.8%**	9.7%	12	74.3	3.8%	11.2%	15	89.3	1.0%	7.6%	23
AP Kone	87.5	**10.2%**	9.8%	15	81.3	**14.5%**	11.2%	10	106.1	**19.5%**	13.6%	7
AP Beckman AU	89.3	**12.1%**	4.0%	14	84.9	**19.4%**	5.7%	16	103.4	**16.4%**	3.9%	14
AP ABX	76.9	−2.7%	10.7%	8	79.6	**11.4%**	16.0%	9	93.7	*5.4%*	7.6%	10
AP Siemens Advia	96.4	**21.4%**	11.7%	2	87.5	**23.3%**	8.3%	5	99.9	**12.8%**	10.4%	11

AP, alkaline-picrate method; IL, Instrumentation Laboratory.
[a] In italics, bias > desirable quality level; in bold, bias > minimum quality level.
[b] Inter-laboratory variability.

Table 3
Results for EQAS materials with creatinine concentrations of about 115 μmol/L for the evaluated years.

Group	Material D 2006 — target value: 122.9 μmol/L ± 2.7 μmol/L				Material C 2010 — target value: 106.1 μmol/L ± 2.3 μmol/L				Material D 2011 — target value: 123.8 μmol/L ± 3.6 μmol/L			
	Mean, μmol/L	Bias[a]	CV[b]	n	Mean, μmol/L	Bias	CV	n	Mean, μmol/L	Bias	CV	n
Enzymatic Roche	114.0	**−7.4%**	5.9%	3	105.2	−1.1%	2.6%	6	122.0	−1.5%	7.4%	9
Enzymatic Others	–	–	–	–	105.1	−1.1%	2.1%	9	121.1	−2.2%	8.3%	14
AP Beckman	129.9	*5.7%*	20.0%	10	107.8	1.7%	10.2%	13	122.0	−1.6%	4.9%	8
AP Biotecnica	124.8	1.6%	14.7%	14	121.1	**13.8%**	6.1%	14	139.7	**12.5%**	6.7%	13
AP Dimension	125.5	2.3%	6.6%	10	109.6	3.3%	6.7%	13	128.2	3.1%	5.7%	16
AP Roche Hitachi	134.4	**9.6%**	3.5%	26	110.5	3.6%	3.9%	30	129.1	*4.1%*	6.7%	32
AP IL	137.9	**12.3%**	7.3%	17	114.0	**7.3%**	6.2%	51	127.3	3.0%	8.1%	51
AP Roche Integra	129.9	**6.1%**	5.8%	15	107.0	0.5%	7.4%	23	124.6	0.8%	6.4%	23
AP Kone	125.5	2.2%	5.0%	15	117.6	**10.7%**	20.0%	11	134.4	**8.2%**	7.3%	7
AP Beckman AU	133.5	**8.9%**	4.8%	15	119.3	**11.9%**	4.0%	22	129.9	*4.7%*	7.8%	21
AP ABX	123.8	0.9%	10.7%	8	110.5	3.7%	15.9%	11	126.4	2.2%	9.8%	11
AP Siemens Advia	142.3	**15.8%**	14.8%	6	116.7	**9.5%**	7.8%	11	131.7	**6.3%**	7.0%	9

AP, alkaline-picrate method; IL, Instrumentation Laboratory.
[a] In italics, bias > desirable quality level; in bold, bias > minimum quality level.
[b] Inter-laboratory variability.

even if the Roche creatinine enzymatic method proved to be efficient and not highly affected by commutability problems [18,19], we cannot exclude that the assigned values may be affected by some non-commutability derived bias.

Moreover the commutability was not checked systematically for all materials and all methods.

4.2. General considerations

Recently, it has been underlined that an effective implementation of standardization in clinical practice requires not just the definition of a whole reference measurement system, including reference measurement procedures, reference materials and accredited reference laboratories, but also the availability of traceable reference intervals and/or decision limits and of an appropriately organized analytical quality control in order to evaluate the performance of commercial systems and of clinical laboratories using them [20]. The reference measurement system for creatinine measurement in serum is now well defined in all its components [5,6], including the definition of appropriate reference intervals for its adoption [21,22]. On the other hand, a well structured EQAS, allowing an adequate and systematic post-market surveillance of the performance of commercial creatinine measuring systems and of laboratories using them, is often lacking [23–25]. With some exceptions, current EQAS are not appropriate to evaluate the traceability of results obtained in clinical laboratories. In fact, the control materials are often not commutable and EQAS too often use consensus-based (i.e., the peer group mean or other indicators for the central tendency) instead of accuracy-based criteria (i.e., target values assigned by the reference procedure) to evaluate the performance of participating laboratories [23].

Finally, acceptability limits for such performance, if any, are usually based on criteria of low hierarchical level, such as goals based on the state of the art of the measurement or set by regulatory bodies, ignoring the impact of measurement performance on the medical use of results.

4.3. Specific comments

Even with some relevant limitations we can draw some conclusions. The presented data indicate a trend of the all method means to converge towards the assigned target values. For creatinine concentrations of about 80 μmol/L for the AP-based methods the % bias reduced from 11.4% in 2006 to 9.0% in 2011, for concentration of about 115 μmol/L from 6.6% to 3.9%, at about 170 μmol/L from 2.9% to 0.1% and at 210 μmol/L from 2.3% to −3.2% (Fig. 1a). Protein concentration, increasing from 47 to 72 g/L, respectively for creatinine concentrations of 80 to 210 μmol/L, being lower at lower creatinine concentration

Table 4
Results for EQAS materials with creatinine concentrations of about 170 μmol/L for the evaluated years.

Group	Material G 2006 — target value: 164.4 μmol/L ± 4.1 μmol/L				Material H 2010 — target value: 174.1 μmol/L ± 5.2 μmol/L				Material H 2011 — target value: 174.1 μmol/L ± 4.0 μmol/L			
	Mean, μmol/L	Bias[a]	CV[b]	n	Mean, μmol/L	Bias	CV	n	Mean, μmol/L	Bias	CV	n
Enzymatic Roche	168.0	2.0%	5.5%	8	168.8	−3.3%	3.7%	8	170.6	−2.3%	6.3%	8
Enzymatic Others	–	–	–	–	168.0	−3.8%	4.8%	8	168.8	−3.0%	4.8%	17
AP Beckman	170.6	3.7%	4.5%	17	171.5	−1.5%	5.2%	17	163.5	**−6.5%**	5.9%	7
AP Biotecnica	161.8	−1.7%	11.6%	14	180.3	3.1%	8.5%	16	180.3	3.2%	6.1%	13
AP Dimension	166.2	0.7%	7.3%	18	175.0	0.5%	5.3%	14	174.1	−0.4%	6.7%	16
AP Roche Hitachi	175.0	**6.2%**	3.8%	33	174.1	0.0%	4.3%	30	175.0	0.4%	3.4%	34
AP IL	172.4	*4.5%*	5.2%	29	174.1	−0.2%	6.4%	52	174.1	−0.4%	6.7%	49
AP Roche Integra	169.7	3.1%	5.6%	18	169.7	−3.0%	5.2%	24	168.0	−3.9%	4.6%	22
AP Kone	163.5	−0.4%	9.2%	15	176.8	1.5%	7.0%	8	177.7	2.0%	8.8%	6
AP Beckman AU	170.6	3.8%	4.8%	19	180.3	3.5%	5.8%	24	177.7	2.0%	4.3%	19
AP ABX	160.9	−2.2%	7.0%	12	169.7	−2.5%	6.7%	11	175.9	1.0%	6.4%	9
AP Siemens Advia	169.7	2.8%	7.6%	9	172.4	−1.2%	10.8%	9	175.9	1.0%	7.3%	11

AP, alkaline-picrate method; IL, Instrumentation Laboratory.
[a] In italics, bias > desirable quality level; in bold, bias > minimum quality level.
[b] Inter-laboratory variability.

Table 5
Results for EQAS materials with creatinine concentrations of about 210 μmol/L for the evaluated years.

Group	Material J 2006 — target value: 210.4 μmol/L ± 5.0 μmol/L				Material I 2010 — target value: 191.8 μmol/L ± 5.5 μmol/L				Material K 2011 — target value: 227.2 μmol/L ± 5.2 μmol/L			
	Mean, μmol/L	Bias[a]	CV[b]	n	Mean, μmol/L	Bias	CV	n	Mean, μmol/L	Bias	CV	n
Enzymatic Roche	206.0	−2.2%	3.2%	8	187.4	−2.3%	2.8%	6	221.0	−2.6%	2.2%	10
Enzymatic Others	–	–	–	–	190.1	−0.9%	3.7%	8	220.1	−3.1%	3.3%	15
AP Beckman	217.5	3.0%	5.8%	15	190.9	−0.4%	6.2%	13	219.2	−3.7%	4.8%	8
AP Biotecnica	205.1	−2.7%	11.1%	12	198.9	3.7%	4.8%	14	215.7	*−5.2%*	8.0%	15
AP Dimension	216.6	2.7%	7.1%	16	197.1	2.9%	3.6%	14	221.9	−2.5%	3.2%	18
AP Roche Hitachi	225.4	**7.2%**	4.4%	34	193.6	1.2%	3.3%	29	225.4	−1.0%	3.9%	34
AP IL	220.1	*4.4%*	4.3%	30	186.5	−2.7%	5.8%	51	218.3	−4.0%	7.2%	51
AP Roche Integra	213.9	1.6%	5.2%	21	185.6	−3.0%	5.5%	23	215.7	*−5.1%*	5.5%	22
AP Kone	204.2	−3.0%	7.1%	15	187.4	−2.3%	9.7%	10	227.2	−0.3%	9.2%	7
AP Beckman AU	217.5	3.4%	4.1%	23	198.0	3.3%	5.8%	23	226.3	−0.4%	6.5%	23
AP ABX	202.4	−4.0%	11.3%	20	185.6	−3.3%	7.1%	11	215.7	*−5.3%*	6.8%	12
AP Siemens Advia	209.5	−0.7%	5.5%	7	186.5	−2.6%	6.4%	9	208.6	**−8.4%**	8.0%	9

AP, alkaline-picrate method; IL, Instrumentation Laboratory.
[a] In italics, bias > desirable quality level; in bold, bias > minimum quality level.
[b] Inter-laboratory variability.

levels, probably did not influence the bias of the AP-based methods. On the contrary it may have attenuated it.

We cannot claim that these biases really represent the performance of the methods on clinical samples, because it is possible that part of this bias could derive from unspecific signal related to the non-commutability of the materials.

But the matrix of the materials remained the same during the years so the data are comparable across time. These moderate improvements are in line with the introduction of a "compensation" factor by some manufacturers of AP-based methods and with the alignment to the IDMS calibration.

So the improvement of the quality of creatinine measurement has started, but, at least in Italy, the effects are reaching the majority or laboratories quite slowly.

The poor analytical specificity of AP-based methods is known from many years [26]. As a result of reaction with plasma pseudo-creatinine chromogens, including proteins, ketones and glucose, methods based on AP reaction may overestimate true serum creatinine concentrations, inducing proportionally greater errors at values lower than 120 μmol/L [27]. This still remains true even after potential elimination of the assay calibration error by alignment to the reference system. In other word, standardization through traceability implementation does not solve the analytical interferences related to an assay's non-specificity [5,27]. The ISO 17511 standard clearly states that, if the reference measurement procedure and corresponding lower-order routine methods have not identical, specificities for the measurand, or at least very similar, traceability cannot be obtained [28].

Trying to solve the problem, some manufacturers marketing AP-based systems have adjusted the calibration by introducing a constant negative offset to "compensate" this error. For example, in the Roche Integra compensated AP assay, a concentration of 18 μmol/L is automatically subtracted from each result [29]. Our data show that this approach may undoubtedly improve the overall alignment of obtained results, especially at creatinine concentrations where the non-specificity effect is more evident (Tables 2 and 3). However, the application of this offset may create a paradoxical effect on higher creatinine concentrations, because of a combined effect with the slope of calibration curve, with a significant underestimation at these still clinically relevant creatinine values (see results of Roche Integra and Siemens Advia groups in Table 5). In general, the use of analytic systems using assays providing lower specificity makes the standardization very difficult or even impossible [30,31]. This issue, together with the high prevalence of laboratories still using these methodologies, may explain why the quality of creatinine determination in Italy is still far from being optimal.

Our picture could have been worsened by some non-commutability phenomena. In fact AP-based methods, being less specific, are more sensitive to possible anomalous components of the matrix of processed materials. A further source of bias could have been introduced by the value assignment performed with an enzymatic method and not with an IDMS method, this fact cannot be excluded completely, but enzymatic methods and particularly Roche enzymatic methods proved their robustness in several occasions [18,19].

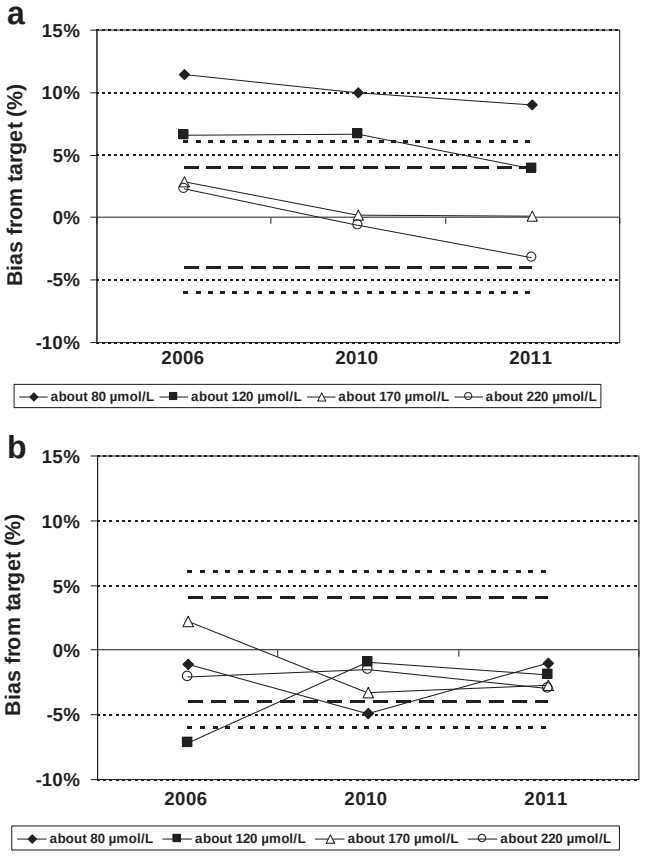

Fig. 1. Percent bias of overall means for the two method macro-categories based on a different analytic principle: AP (a) and enzymatic (b). The dotted and the dashed line indicate the maximum acceptable bias at desirable (±4.0%) and at minimum quality level (±6.0%), respectively.

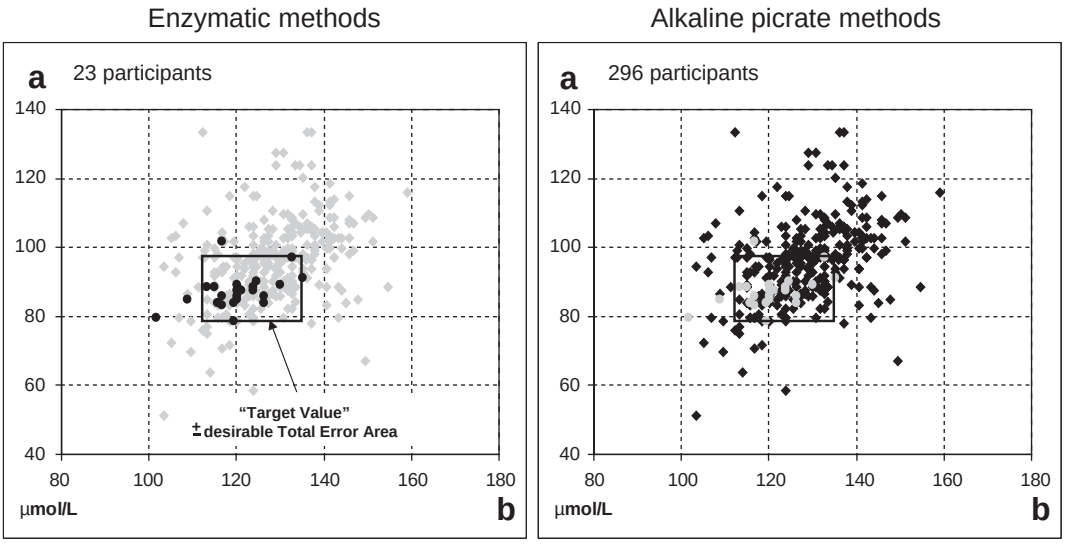

Fig. 2. Youden plots for creatinine results of control materials B 2011 (a) and D 2011 (b) by individual laboratories participating in EQAS Prolarit.

These facts could have contributed to the higher dispersion of the AP-based methods results shown in Fig. 2, however just the high interlaboratory variability within a specific method group (frequently higher than 10%, see Tables 2 and 3) is sufficient to produce a substantial proportion of data outside the ±8.9% TE.

Enzymatic methods are not completely free of defects, but largely eliminate analytical non-specificity and their superiority has been proved and constantly confirmed [18,30–32]. In addition, their use can also significantly improve the precision of creatinine measurement [27,33]. In our study, the evaluation of the variability of results within different groups confirmed a lower inter-laboratory CV for the enzymatic groups. It is particularly relevant that, even though the "Enzymatic Others" group was quite heterogeneous in the composition [in 2011 several different manufacturers were included: Instrumentation Laboratory (n = 7), Beckman Coulter (n = 5), Siemens Healthcare Diagnostics (n = 4) and 2 Abbott (n = 2)], its reported CVs were lower than those of more homogeneous AP-based groups. We should, however, highlight that enzymatic assays were also not perfectly aligned to our reference, showing during 2011 an average bias of −2.1%, which was very stable at different tested concentrations. This implies that also standardization of enzymatic assays can further be improved. As a positive signal, the number of laboratories participating to EQAS Prolarit using enzymatic methods is increasing: in 2006 only 4 laboratories (~1%) were using enzymatic methods, while in 2011 they were 25 (7.0% of the total).

5. Conclusions

The main aim of this study was to evaluate the impact of the recent standardization process on the quality of creatinine measurement in Italy. With few exceptions, the obtained results did not show significant differences between 2006 and the most recent years. Compared to previous reports [10], some improvement can be noticed, especially for some groups of systems and for intermediate concentration values. The wide use of AP-based methods represents today a major obstacle to the achievement of the analytical quality needed for the optimal clinical use of such an important test. The evidence in favor of the adoption of enzymatic assays, presenting the necessary analytical specificity, is undeniable. As already said [27], it is our belief that AP-based assays should be abandoned. Taking this step by our profession becomes now not only a technological duty, but also an ethical obligation [34].

Acknowledgments

The Authors wish to thank laboratories participating to EQAS Prolarit (http://www.prolarit.it/default.asp). The majority of results reported in this paper were previously published in an Italian journal (Biochim Clin 2012;36:414) to promote their knowledge at the national level.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.cca.2013.10.001.

References

[1] National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Kidney Disease Outcome Quality Initiative. Am J Kidney Dis 2002;39:S1–S266.
[2] Levey AS, Bosch JP, Lewis J, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Ann Intern Med 1999;130:461–70.
[3] 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.
[4] Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro 3rd AF, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med 2009;150:604–12.
[5] Panteghini M. Estimating glomerular filtration rate from serum creatinine measurements: analytical issues and standardization programs. RIMeL/IJLaM 2007;3:56–60 [Suppl.].
[6] Panteghini M, Myers GL, Miller WG, Greenberg N. The importance of metrological traceability on the validity of creatinine measurement as an index of renal function. Clin Chem Lab Med 2006;44:1287–92.
[7] Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices. Off J Eur Communities Dec 7 1998; L331:1–37.
[8] Dati F. The new European directive on in vitro diagnostics. Clin Chem Lab Med 2003;41:1289–98.
[9] Delanghe JR, Cobbaert C, Galteau MM, Harmoinen A, Jansen R, Kruse R, 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.
[10] Ceriotti F, Infusino I, Luraschi P, Panteghini M. Evaluation of the trueness of serum creatinine measurement: results of a group of Italian laboratories. Biochim Clin 2007;31:19–23.
[11] Miller WG, Jones GRD, Horowitz GL, Weykamp C. Proficiency testing/external quality assessment: current challenges and future directions. Clin Chem 2011;57:1670–80.

[12] CLSI. Characterization and qualification of commutable reference materials for laboratory medicine; approved guideline. CLSI Document C53-A. Wayne, PA: Clinical and Laboratory Standards Institute; 2010.
[13] 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.
[14] Infusino I, Luraschi P, Valente C, Panteghini M. Accuracy improvement of creatinine measurement by use of an enzymatic assay. Clin Chem Lab Med 2007;45:A155.
[15] http://www.westgard.com/biodatabase1.htm . [Accessed October 2012].
[16] Fraser CG, Hyltoft Petersen P, Libeer JC, Ricos C. Proposal for setting generally applicable quality goals solely based on biology. Ann Clin Biochem 1997;34:8–12.
[17] Stöckl D, Reinauer H. Development of criteria for the evaluation of reference method values. Scand J Clin Lab Invest 1993;212:16–8.
[18] Greenberg N, Roberts WL, Bachmann LM, Wright EC, Dalton RN, Zakowski JJ, et al. Specificity characteristics of 7 commercial creatinine measurement procedures by enzymatic and Jaffe method principles. Clin Chem 2012;58:391–401.
[19] Miller WG, Myers GL, Ashwood ER, Killeen AA, Wang E, Thienpont LM, et al. Creatinine measurement. State of the art in accuracy and interlaboratory harmonization. Arch Pathol Lab Med 2005;129:297–304.
[20] Panteghini M. Implementation of standardization in clinical practice: not always an easy task. Clin Chem Lab Med 2012;50:1237–41.
[21] Ceriotti F, Boyd JC, Klein G, Henny J, Queraltó J, Kairisto V. Panteghini M; on behalf of the IFCC Committee on Reference Intervals Decision Limits (C-RIDL). Reference intervals for serum creatinine concentrations: assessment of available data for global application. Clin Chem 2008;54:559–66.
[22] Panteghini M, Ceriotti F. Obtaining reference intervals traceable to reference measurement systems: it is possible, who is responsible, what is the strategy? Clin Chem Lab Med 2012;50:813–7.
[23] 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.
[24] Myers GL, Miller WG, Coresh J, Fleming J, Greenberg N, Greene T, 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.
[25] Friedecky B, Kratochvila J, Budina M, Jabor A. Quality of serum creatinine measurement in light of EQA programs. Clin Chem Lab Med 2007;45:685–8.
[26] Spencer K. Analytical reviews in clinical biochemistry: the estimation of creatinine. Ann Clin Biochem 1986;23:1–25.
[27] Panteghini M, on behalf of the IFCC Scientific Division. Enzymatic assays for creatinine: time for action. Clin Chem Lab Med 2008;46:567–72.
[28] ISO 17511. In vitro diagnostic medical devices — measurement of quantities in biological samples — metrological traceability of values assigned to calibrators and control materials. Geneva: ISO; 2003.
[29] Junge W, Wilke B, Halabi A, Klein G. Determination of reference intervals for serum creatinine, creatinine excretion and creatinine clearance with an enzymatic and a modified Jaffé method. Clin Chim Acta 2004;344:137–48.
[30] Drion I, Cobbaert C, Groenier KH, Weykamp C, Bilo HJB, Wetzels JFM, et al. Clinical evaluation of analytical variations in serum creatinine measurements: why laboratories should abandon Jaffe techniques. BMC Nephrol 2012;13:133.
[31] Boutten A, Bargnoux AS, Carlier MC, Delanaye P, Rozet E, Delatour V, et al. Enzymatic but not compensated Jaffe methods reach the desirable specification of NKDEP at normal levels of creatinine. Results of the French multicentric evaluation. Clin Chim Acta 2013;419:132–5.
[32] Cobbaert CM, Baadenhuijsen H, Weykamp CW. Prime time for enzymatic creatinine methods in pediatrics. Clin Chem 2009;55:549–58.
[33] Infusino I, Luraschi P, Valente C, Panteghini M. Overall improvement in imprecision of serum creatinine measurement using an enzymatic assay. Clin Chem Lab Med 2007;45:S275.
[34] Bossuyt X, Louche C, Wiik A. Standardisation in clinical laboratory medicine: an ethical reflection. Ann Rheum Dis 2008;67:1061–3.