Con: Should we abandon the use of the MDRD equation in favour of the CKD-EPI equation?

Pierre Delanaye¹, Hans Pottel² and Rossini Botev³
¹Department of Nephrology-Dialysis-Transplantation, University of Liège, CHU Sart Tilman, Liège, Belgium,
²Interdisciplinary Research Center, University of Leuven, Kortrijk, Belgium and
³Department of Nephrology, Hawaii Permanente Medical Group, Honolulu, HI, USA

INTRODUCTION

The best overall index of renal function is the glomerular filtration rate (GFR) [1]. Since measuring the GFR can be cumbersome and costly, estimation of GFR is essential for the diagnosis and evaluation of chronic kidney disease (CKD), defined as kidney damage or GFR <60 mL/min/1.73 m² for ≥3 months and staged by levels of GFR according to the NKF KDOQI (The National Kidney Foundation Kidney Disease Outcomes Quality Initiative) guidelines published in 2002 [2]. These now decade-old guidelines are under consideration for revision by the KDIGO (Kidney Disease: Improving Global Outcomes) CKD Work Group [3]. Several authors have proposed in the past different equations to estimate GFR based on serum creatinine concentration, the latter being determined by the individual's muscular mass in steady state, and thus by age, gender and ethnicity [4-8]. Among these equations, the one proposed by Cockcroft and Gault in 1976 [4] was doubtless the most popular for many years because of its simplicity and ease to calculate at the patient's bedside. However, this equation is an estimation of the creatinine clearance, not of the GFR, and was particularly imprecise, notably in CKD patients. In 1999, an equation derived from measured GFR (mGFR) was developed in the Modification of Diet in Renal Disease (MDRD) Study. This equation included the serum creatinine concentration and six other variables, subsequently abbreviated to four variables [5, 6]. This equation has been endorsed by the KDOQI guidelines and is currently used by most clinical laboratories [2]. Several limitations of the MDRD study equation were elaborated on in the literature [9-12], the major one being the systematic underestimation of mGFR >60 mL/min/1.73 m², which essentially translates in to overestimation of the CKD prevalence. A new equation, the CKD-Epidemiology Collaboration (CKD-EPI) equation, was proposed in 2009 for better estimation of GFR levels >60 mL/min/1.73 m² [7].

ASSESSING THE PERFORMANCE OF EQUATIONS (OR WHY WE NEED PRECISION)

Four main statistical tools are used when performance of equations must be evaluated: bias, precision, accuracy and CKD classification.

Bias is defined as the mean (or median) difference between the estimated GFR (eGFR) and mGFR [13, 14]. It is a conceptualization of the systematic error. Because this error is systematic, it is relatively easy to assess. For example, if a particular equation systematically underestimates the mGFR by 5 or 10 mL/min/1.73 m², it is not difficult for the clinicians to interpret the eGFR result for a given patient. Conversely, when considering the eGFR results in population studies, such a systematic error will be misleading in terms of CKD prevalence or CKD-related mortality and morbidity-associated risk [15, 16]. This was specifically observed with the MDRD equation, which by underestimating the mean mGFR, overestimates the CKD stage 3 prevalence in the general population [7, 15-18].

Precision is a concept that is frequently forgotten by most authors but it is, however, fundamental. Precision is usually expressed as standard deviation (SD) or interquartile range (IQR) of the bias, and thus represents the spread of 68% (assuming normal distribution) or 50% of the values around the bias, respectively. It conceptualizes the random error, which is much more difficult to assess and correct than the bias. The conceptualization is easier for the clinician to understand when precision is expressed as SD instead of IQR, because the former permits the estimation of the 95% range of eGFR values. Contrary to the bias, this type of error will have a low impact on the results in a population, e.g. the CKD prevalence. However, such an error could be misleading for the clinicians when interpreting an individual's eGFR result. In a patient with an mGFR of 60 mL/min/1.73 m², an equation with a perfect bias (0 mL/min/1.73 m²) but a global precision (SD) of 10 mL/min/1.73 m² would provide 95% probability so that the eGFR value is between 40 and 80 mL/min/1.73 m²,
something the clinicians would not be aware of when an eGFR value is routinely reported by a clinical laboratory.

**Accuracy** combines bias and precision and it is easy to understand when presented as the percentage of eGFRs within a defined range of their respective mGFRs. Root mean square error also combines bias and precision but its conceptualization is more difficult to interpret in practice by clinicians. In the KDOQI guidelines, a range of ±30% with 90% of eGFRs in this range, was cited as a useful measure of accuracy [2]. As practicing clinicians, we have to question if such a goal is really sufficient for important clinical decisions, as, for example, whether a potential candidate could be a living kidney donor. In such situations, an accuracy of ±15% could be more useful [9, 14].

**CKD classification** is one of the most important statistical tools for the clinicians when evaluating the performance of an equation on an individual level [9, 19]. It is related to the percentage of patients correctly classified by eGFR into the different CKD stages in comparison with the confirmatory test of mGFR [20]. Unfortunately, this statistical tool is not always used appropriately [9]. For example, some authors [7, 21] consider the eGFR as a reference, while many others correctly used the mGFR as a reference for defining the five CKD stages when assessing an equation’s performance [19, 22-32]. An equation with high correct CKD classification would decrease the need for determining the mGFR and provide great confidence to the clinicians to implement an appropriate plan of action according to the individual’s eGFR result.

**THE CKD-EPI EQUATION: WHAT IS (REALLY) NEW?**

The relationship between GFR and serum creatinine is different in healthy subjects and CKD patients [33, 34]. This physiological fact explains why the MDRD equation, developed exclusively from a CKD population, underestimates the high GFR levels. Because the authors of the CKD-EPI equation included a significant proportion of subjects with a ‘normal’ GFR, the serum creatinine variable was modelled as a spline with sex-specific knots (0.7 and 0.9 mg/dL for women and men, respectively) and different exponents were applied to serum creatinine according to its level [7]. As expected, the performance of the new equation was globally better than the MDRD one. For example, in the external validation data set, the median bias (eGFR-mGFR), precision (IQR) and accuracy ±30% for the MDRD and CKD-EPI equations were: -5.5 versus -2.5 mL/min/1.73 m², 18.3 versus 16.6 mL/min/1.73 m² and 81 versus 84%, respectively. However, it must be underlined that the improved performance of the CKD-EPI equation was essentially due to a lower bias [3 mL/min/1.73 m² (55%) reduction], while precision remained comparable [1.7 mL/min/1.73 m² (9%) reduction] and suboptimal. Accuracy ±30% was essentially unchanged [3.5% (4%) increase]. As expected too, for GFR >60 mL/min/1.73 m² the CKD-EPI equation's bias was better [7 mL/min/1.73 m² (67%) reduction], but meanwhile its precision declined by 12.9 mL/min/1.73 m² (114%), 11.3 versus 24.2 mL/min/1.73 m² for mGFR <60 and >60 mL/min/1.73 m², respectively, and was not significantly different compared with the MDRD equation [1.5 mL/min/1.73 m² (6%) reduction] [7].

How can we explain this absence of a better precision in the new equation? Two main explanations could be advanced. First, the precision of an estimator will strongly depend on the precision of measuring serum creatinine (as a main biological variable) and GFR. In the vast majority of studies used as development data set for the CKD-EPI equation, serum creatinine has been measured by the Jaffe methods which are largely less precise, even if isotope dilution mass spectrometry traceable, than the enzymatic methods. This imprecision impacts the global accuracy of the new equation [35-37]. Hence, it is not surprising that the best accuracy of creatinine-based equations is seen in studies with serum creatinine measured by enzymatic assay in paediatric [38] or adult [14, 39] studies. In the development data set, only urinary clearances of subcu-taneously injected iothalamate were performed to measure GFR. In this measurement, precision is not the best [40] and, incidentally, the plasma clearance of iohexol has been shown to have a better precision and reproducibility [41-43]. Secondly, the relative lack of precision of the CKD-EPI equation is probably due to the studied population [7]. Certainly, the population sample is impressive as the new equation has been built from a development data set including 5504 subjects (with an internal validation data set of 2750 subjects) and subsequently tested in an external validation data set of 3896 subjects. Although the MDRD cohort represented a relatively homogeneous population (CKD patients without diabetes), the CKD-EPI one included a more heterogeneous population with potentially different GFR-serum creatinine relationships [33, 34].

After its introduction, the CKD-EPI equation was studied in various populations. A PubMed database (last accessed 18 June 2012) search for GFR, MDRD and CKD-EPI in adults with a minimum of 50 mGFRs (estimates from smaller studies can be unreliable) and provided data for ±30% accuracy recovered 26 publications [7, 21, 26, 32, 39, 44-64]. The results for accuracy, bias and precision and their respective calculated
weighted average values are presented in Table 1. The CKD-EPI equation had a slightly better weighted average for ±30% accuracy of 1.8%, mean/median bias of 3.5/0.1 mL/min/1.73 m², respectively, and precision of 1.1 mL/min/1.73 m² compared with the MDRD one. These differences are not clinically significant with the exception of a better mean bias of 3.5 mL/min/1.73 m² when considering the very low GFR levels. The differences between the two equations by weighted average analysis for strata of GFR >60 mL/min/1.73 m² were again clinically insignificant.

CONCLUSION: POPULATION VERSUS PATIENT IMPLICATIONS

At best, the CKD-EPI equation might be considered as an evolution but not a revolution. By improving the systematic error of the CKD-EPI equation in comparison with the MDRD one, the advantage of the new equation is when we think in terms of population [7, 14, 17, 65]. In this view, it seems that the CKD-EPI equation performs better to define the CKD stage 3 prevalence in a general population, and especially in the younger subjects [66]. However, without the confirmatory mGFR test, this conclusion is not reliable and there are still reasons to think that the new equation overestimates the CKD prevalence in the Caucasian population [17]. In the same thought, it has been recently proved in different cohorts that the better bias linked to the CKD-EPI equation leads to a better prediction of the CKD-associated risk of mortality [65, 67, 68]. Although impressive, we have already underlined the limitations of this type of epidemiological studies [69].

The superiority of the CKD-EPI equation could be accepted in epidemiological trials but this is certainly more questionable if we are thinking in terms of 'the patient'. If we have to know and follow the GFR of a given individual, the concept of precision becomes very important and in this context, the CKD-EPI equation is clearly lacking any additional value. This assertion is reinforced if we consider the percentage of subjects who are correctly classified into the five CKD stages. In the validation data set, Levey et al. [7] demonstrated a concordance between the measured and estimated CKD stages of 69%, while Murata et al. [21] observed a correct classification of >70% only in 6 of the 25 GFR studied groups. Bjork et al. [32] found an overall proper CKD classification of 69%. All these results should be considered suboptimal as in the case of the 38% overall CKD misclassifications reported for the MDRD equation in a recent review [9].

Obviously, both the CKD-EPI and MDRD equations are not 'magic'. Serum creatinine remains the principal variable of the GFR-estimating equations, and if it is not accurately representing the individual's muscular mass [70], there is little chance that including the creatinine into the equations would improve their performance. Classical examples are hyperfiltrating diabetic patients [49, 59, 71] and cirrhotic [72, 73], renal transplanted [21, 51, 55], elderly [21, 73] and anorectic patients [73].

As clinicians, we have to know the GFR of individuals, not of populations. We have to treat patients, not statistical risks by associations. Therefore, we have to minimize the random error of the estimator, not only the systematic one. In fact, the true question might not be 'Is the CKD-EPI equation better for the estimation of my given patient?' but maybe 'Is there any chance that any creatinine-based equation is precise enough in my patient?' and 'Would it not be better to measure GFR by a reference method in my specific patient?' [20].
### Table 1. Performance of MDRD and CKD-EPI equations (using calibrated SCr values) in studies reporting accuracy data for both equations and with minimum of 50 measured GFRs

<table>
<thead>
<tr>
<th>Study</th>
<th>GFR method</th>
<th>SCr calibration</th>
<th>Population</th>
<th>N GFRs</th>
<th>Mean mGFR ± SD (range)</th>
<th>Accuracy 30%</th>
<th>Bias 15%</th>
<th>Precision</th>
<th>Study</th>
<th>GFR method</th>
<th>SCr calibration</th>
<th>Population</th>
<th>N GFRs</th>
<th>Mean mGFR ± SD (range)</th>
<th>Accuracy 30%</th>
<th>Bias 15%</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Murata et al. [21]</td>
<td>Iothalamate</td>
<td>Yes IDMS</td>
<td>Mixed</td>
<td>5238</td>
<td>56 ± 30</td>
<td>77.6</td>
<td>78.4</td>
<td>-4.1</td>
<td>Levey et al. [7]</td>
<td>Iothalamate</td>
<td>Yes IDMS</td>
<td>Mixed</td>
<td>3896</td>
<td>68 ± 36</td>
<td>80.6</td>
<td>84.1</td>
<td>5.5</td>
</tr>
<tr>
<td>Eriksen et al. [39]</td>
<td>Iothalamate</td>
<td>Yes IDMS</td>
<td>General (no CKD)</td>
<td>1621</td>
<td>92 ± 14</td>
<td>93.9</td>
<td>95.0</td>
<td>1.3</td>
<td>Bjork et al. [32]</td>
<td>Iothalamate</td>
<td>Yes IDMS</td>
<td>Mixed</td>
<td>1397</td>
<td>44 (12-116)</td>
<td>79.5</td>
<td>79.1</td>
<td>-2.0</td>
</tr>
<tr>
<td>Buron et al. [58]</td>
<td>Inulin</td>
<td>Yes LCMS</td>
<td>KT recipients</td>
<td>1249</td>
<td>54 ± 18 (15-90)</td>
<td>85.0</td>
<td>81.0</td>
<td>-0.5</td>
<td>Buron et al. [58]</td>
<td>Inulin</td>
<td>Yes LCMS</td>
<td>KT recipients</td>
<td>1249</td>
<td>54 ± 18 (15-90)</td>
<td>85.0</td>
<td>81.0</td>
<td>-0.5</td>
</tr>
<tr>
<td>Nyman et al. [47]</td>
<td>Iothalamate</td>
<td>Yes IDMS</td>
<td>Mixed</td>
<td>850</td>
<td>55 (9-121)</td>
<td>79.9</td>
<td>79.5</td>
<td>1.0</td>
<td>Iliadis et al. [57]</td>
<td>Iothalamate</td>
<td>Yes IDMS</td>
<td>Mixed</td>
<td>448</td>
<td>73 ± 23</td>
<td>78.8</td>
<td>80.7</td>
<td>7.5</td>
</tr>
<tr>
<td>Lane et al. [60]</td>
<td>Iothalamate</td>
<td>Yes CI Clin</td>
<td>Pre- and post nephrectomy</td>
<td>425</td>
<td>50 (median) (4-142)</td>
<td>75</td>
<td>80.0</td>
<td>-1.0</td>
<td>Lane et al. [60]</td>
<td>Iothalamate</td>
<td>Yes CI Clin</td>
<td>Pre- and post nephrectomy</td>
<td>425</td>
<td>50 (median) (4-142)</td>
<td>75</td>
<td>80.0</td>
<td>-1.0</td>
</tr>
<tr>
<td>Cirillo et al. [56]</td>
<td>Iothalamate</td>
<td>Yes IDMS</td>
<td>Mixed</td>
<td>356</td>
<td>72 ± 36</td>
<td>87.4</td>
<td>88.2</td>
<td>-5.2</td>
<td>Teo et al. [54]</td>
<td>99mTc-DTPA</td>
<td>Yes IDMS</td>
<td>CKD</td>
<td>232</td>
<td>52 ± 28</td>
<td>79.7</td>
<td>82.8</td>
<td>-1.0</td>
</tr>
<tr>
<td>Michels et al. [26]*</td>
<td>Iothalamate</td>
<td>Yes IDMS</td>
<td>Mixed</td>
<td>271</td>
<td>73 ± 30</td>
<td>81.2</td>
<td>84.5</td>
<td>0.8</td>
<td>Michels et al. [26]*</td>
<td>99mTc-DTPA</td>
<td>Yes IDMS</td>
<td>CKD</td>
<td>271</td>
<td>73 ± 30</td>
<td>81.2</td>
<td>84.5</td>
<td>0.8</td>
</tr>
<tr>
<td>Teo et al. [54]</td>
<td>Iothalamate</td>
<td>Yes IDMS</td>
<td>CKD</td>
<td>232</td>
<td>52 ± 28</td>
<td>79.7</td>
<td>82.8</td>
<td>-1.0</td>
<td>Teo et al. [54]</td>
<td>99mTc-DTPA</td>
<td>Yes IDMS</td>
<td>CKD</td>
<td>232</td>
<td>52 ± 28</td>
<td>79.7</td>
<td>82.8</td>
<td>-1.0</td>
</tr>
<tr>
<td>White et al. [46]</td>
<td>Iothalamate</td>
<td>Yes IDMS</td>
<td>KT recipients</td>
<td>207</td>
<td>58 ± 22</td>
<td>79</td>
<td>84.0</td>
<td>-8.0</td>
<td>White et al. [46]</td>
<td>99mTc-DTPA</td>
<td>Yes IDMS</td>
<td>KT recipients</td>
<td>207</td>
<td>58 ± 22</td>
<td>79</td>
<td>84.0</td>
<td>-8.0</td>
</tr>
<tr>
<td>Redal-Baigorri et al.  [48]*</td>
<td>Iothalamate</td>
<td>Yes IDMS</td>
<td>Oncology</td>
<td>185</td>
<td>85 ± 20</td>
<td>88.6</td>
<td>89.7</td>
<td>0.8</td>
<td>Redal-Baigorri et al.  [48]*</td>
<td>99mTc-DTPA</td>
<td>Yes IDMS</td>
<td>Oncology</td>
<td>185</td>
<td>85 ± 20</td>
<td>88.6</td>
<td>89.7</td>
<td>0.8</td>
</tr>
<tr>
<td>Poge et al. [55]</td>
<td>Iothalamate</td>
<td>Yes IDMS</td>
<td>KT recipients</td>
<td>170</td>
<td>40 (12-83)</td>
<td>71.8</td>
<td>64.1</td>
<td>4.5</td>
<td>Poge et al. [55]</td>
<td>99mTc-DTPA</td>
<td>Yes IDMS</td>
<td>KT recipients</td>
<td>170</td>
<td>40 (12-83)</td>
<td>71.8</td>
<td>64.1</td>
<td>4.5</td>
</tr>
<tr>
<td>Study</td>
<td>Methodology</td>
<td>Participants</td>
<td>Study Design</td>
<td>GFR Estimation</td>
<td>Creatinine</td>
<td>mGFR</td>
<td>±15%</td>
<td>±30%</td>
<td>Mean Bias</td>
<td>Median Bias</td>
<td>Bias</td>
<td>Relative Accuracy</td>
<td>Kukla et al. [51]</td>
<td>MDRD, simplified (4 variables), re-expressed with calibrated serum creatinine, Modification of Diet in Renal Disease Study equation; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration equation; GFR, glomerular filtration rate; SCr, serum creatinine calibrated to IDMS (isotope dilution mass spectroscopy), LCMS (liquid chromatography-mass spectrometry) or CIClin (Cleveland Clinic Laboratory); mGFR, measured GFR in mL/min/1.73 m²—urinary clearance unless otherwise described; Accuracy, % of GFR estimates within ±15% and ±30% of mGFRs; Bias, estimated minus measured GFR in mL/min/1.73 m² (values for the mean bias in italic type were calculated if possible whenever not provided); Precision, one standard deviation (SD) of mean bias in mL/min/1.73 m²; 99mTc-DTPA, 99mtechnetium-diethylenetriamine pentaacetic acid; 51Cr-EDTA, 51chromium-ethylene diamine tetracetic acid; CKD, chronic kidney disease; KT, kidney transplant; DM, diabetes mellitus.</td>
<td>Study design included some SCr measurements not exactly on the day of GFR measurement.</td>
<td>Data calculated for accuracy ±30%, accuracy ±15%, mean bias, median bias and precision from analysis of 18245, 690, 12303, 9484 and 3572 mGFRs, respectively.</td>
<td>Data calculated for accuracy ±30%, accuracy ±15%, mean bias, median bias and precision from analysis of 1950, 132, 1072, 1048 and 1000 mGFRs, respectively.</td>
</tr>
</tbody>
</table>
CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES


17. Delanaye P, Cavalier E, Mariat C et al. MDRD or CKD-EPI study equations for estimating prevalence of stage 3 CKD in epidemiological studies: which difference? Is this difference relevant? BMC Nephrol 2010; 11: 8


Diabetes Care 2007; 30: 300-305


epidemiology collaboration equation any better? Clin Chem 2010; 56: 474-477


49. Camargo EG, Soares AA, Detanico AB et al. The Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation is less accurate in patients with Type 2 diabetes when compared with healthy individuals. Diabet Med 2011; 28: 90-95


OPPONENT’S COMMENTS

We congratulate Dr Delanaye and colleagues on their scholarly review of GFR estimation and the comparative performance of the CKD-EPI versus MDRD Study equations. We agree with them about the importance of ascertainment of GFR in clinical decisions, the practicality of estimating the GFR from serum creatinine, the greater accuracy of the CKD-EPI versus MDRD study equation and the persisting imprecision of GFR estimates as an irremediable limitation of variation in non-GFR determinants of serum creatinine. We also agree that the CKD-EPI equation represents an evolution, not a revolution, in GFR estimation, but we disagree with their conclusion that we should not make the change to report the eGFR using the CKD-EPI equation. At present, there are hundreds of millions of creatinine measurements performed each year for the purpose of estimating the GFR. For the vast majority of these measurements, it is not possible to measure GFR.

Our disagreement rests on three main arguments. First, their discussion on the limitations in measurement methods is not relevant when comparing the performance of two equations in the same dataset. Well-done studies, such as ours [1], show a consistent improvement in bias at a higher GFR and in overall accuracy [2]. Second, they erroneously characterize the performance of the two equations as comparable. While we agree that the improvement in precision is small, with the improvement in bias, the resulting improvement in accuracy is clinically important. For example, in our study, the percent of large errors (>30% of the measured GFR) decreased from 19.4 to 15.9% (18%) across the range of GFRs, and from 17.7 to 13.2% (25%) for eGFR 60-89 ml/min per 1.73 m² [1]. These results represent a large and meaningful improvement in performance. Third, they dismiss the reduction in prevalence estimates and improvement in risk stratification as not relevant to clinical practice. However, these improvements translate directly to more accurate individual decision making for disease detection and prognosis, key elements in patient care. Moreover, these improvements are observed in large subgroups at low risk in whom there has been concerns over overdiagnosis of CKD (middle-aged people, women and whites). Reporting eGFR using the CKD-EPI equation would mean that substantially fewer patients would have to worry about an erroneous diagnosis of kidney disease and the risk for the future.

At this time, the CKD-EPI equation maximizes the information available from the serum creatinine concentration. Evolution is a slow process. Gradual changes cannot be ignored.

Lesley A. Inker and Andrew S. Levey

REFERENCES
