The Clinician and Estimation of Glomerular Filtration Rate by Creatinine-based Formulas: Current Limitations and Quo Vadis

Rossini Botev,* Jean-Pierre Mallié,† Jack F.M. Wetzels,* Cécile Couchoud,* and Otto Schäck‡

Summary

The GFR has a paramount diagnostic and staging role in the Kidney Disease Outcome Quality Initiative Clinical Practice Guidelines for Chronic Kidney Disease (K/DOQI-CKD). The most widely used serum creatinine-based formulas for estimated GFR (eGFR) are the Cockcroft-Gault (CG) and Modification of Diet in Renal Disease Study (MDRD). Recently, a new Chronic Kidney Disease Epidemiology Collaboration equation has been developed. Review of the literature revealed that CG and MDRD formulas correctly assigned overall only 64% and 62%, respectively, of the subjects to their actual K/DOQI-CKD classification’s GFR groups as determined by measured GFR (mGFR). This suggests that approximately 10 million (38%) subjects may have been misclassified on the basis of estimated CKD prevalence of 26.3 million adults in the United States. The purpose of this review is to help the clinician understand the limitations of using eGFR in daily practice. We also elaborate upon issues such as the differences among markers of mGFR, the validity of adjusting GFR for body surface area in certain populations, the limited data on boundaries for normal mGFR according to age, gender, and race, the need for calibration of a wide spectrum of serum creatinine measurements, the lack of actual eGFR value above 60 ml/min per 1.73 m² and reference for normal mGFR in the clinical laboratories’ reports, and the performance evaluation of the eGFR formulas. Several pitfalls have to be overcome before we can reliably determine health and disease in daily nephrology practice to preserve the first rule of practicing medicine: primum non nocere.


Introduction

The best overall index of renal function is considered to be the glomerular filtration rate (GFR) (1,2). In the current Kidney Disease Outcome Quality Initiative Clinical Practice Guidelines for Chronic Kidney Disease (K/DOQI-CKD), all individuals with GFR <60 ml/min per 1.73 m² for ≥3 months are classified with chronic kidney disease (CKD), and different GFR levels are used for CKD staging, where “inulin clearance is widely regarded as the gold standard for measuring GFR” (2). As an alternative to the cumbersome and costly process of GFR measurement, several serum creatinine (SCr)-based formulas for estimating the renal function in adults were created in the past (3), and the most widely used ones are the Cockcroft-Gault (CG) formula (4), which predicts urinary creatinine clearance, and the Modification of Diet in Renal Disease Study (MDRD) prediction equation (5) for GFR. We previously described in detail the differences between these two formulas and their performance in comparison with urinary inulin clearance (Cin) (6).

The K/DOQI-CKD endorsed the MDRD equation as a clinically useful estimate of GFR and on the basis of this equation, the estimated CKD prevalence in multiple countries ranges from 10% to 15% (7–10). In the United States, the estimated CKD (stages 1 to 4) prevalence according to both albuminuria and decreased estimated GFR (eGFR) was 13.1%, representing 26.3 million people on the basis of the 2000 United States population census (7). The estimated prevalence of CKD in any given population depends on the formula used for eGFR because applying the recently created Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation reduced the estimated prevalence from 13.1% to 11.5% (11), representing a decrease of 3.2 million people. Nevertheless, CKD still remains a significant public health issue.

Therefore, a proper estimation of the GFR has a pivotal role for the diagnosis and staging of CKD and is essential in nephrology practice. Furthermore, the eGFR has a very important role in large epidemiologic studies whenever GFR measurement is not feasible for practical and cost-containing reasons. The K/DOQI-CKD acknowledged that eGFR methods “are not completely sensitive or specific in detecting decreased GFR” and “thus, misclassification is possible, and clinicians should carefully consider all aspects of the patient’s clinical presentation in interpreting test results and determining evaluation and management.” Usually, an eGFR formula performance is assessed by the bias, precision, and accuracy of the eGFR values in comparison with the respective measured GFR (mGFR) ones. The role of mGFR as a confirmatory
Measurement of GFR

Measurement of GFR is based on clearance of a substance that is not bound to proteins, freely filtered by the glomeruli, and it is not synthesized, transported, or metabolized by the kidneys. Creatinine meets most, but not all, of the requirements for an ideal endogenous GFR marker. It is secreted and reabsorbed by the renal tubules and extrarenally eliminated in severe CKD (13). These non-GFR determinants of urinary and serum concentrations introduce significant errors when calculating the creatinine clearance as a GFR measurement (2). Furthermore, the tubular creatinine secretion varies at different levels of mGFR (14) and in patients with severe proteinuria (15).

Cin is accepted as the gold standard for mGFR (2,16,17). Several experimental studies by direct nephron micropunctures have confirmed that inulin fulfills the requirements for an ideal filtration marker (18–21). Such studies cannot be practically conducted in humans for ethical reasons. In steady-state conditions, inulin is administered by continuous intravenous infusion to achieve constant serum concentration, and several timed urine collections are used to calculate the GFR. Unfortunately, the Cin method cannot be used routinely in daily practice because of its complexity. Other methods for mGFR using urinary or plasma clearance of exogenous agents such as iothalamate, ethylene diaminetetraacetic acid (EDTA), diethylene-triamine pentaacetic acid (DTPA), or iohexol is beyond the scope of this review exploring differences between Cin and other GFR markers as EDTA, DTPA, or iohexol is beyond the scope of this review and was recently discussed (12).

In summary, from these small studies, it is evident that Cio may differ significantly from Cin at different Cin levels, and thus a eGFR formula based on Cio as mGFR might have different discriminatory ability to properly estimate the GFR. Previous studies used mainly the correlation coefficient to assess the differences between Cio and Cin, but this does not reveal the dispersion of Cio values, and the best method to apply is accuracy. Obviously, there is a need to standardize the methodology for GFR measure-
ment and establish calibration coefficients for the currently used GFR markers by simultaneous measurements in much larger and diverse populations.

GFR Adjustment for Body Surface Area (BSA)

In the 19th century, it was accepted that metabolic rate is proportional to BSA, and indexing physiologic variables, including GFR, to BSA was introduced to adjust for differences in body size. Delanaye et al. (33,34) and Geddes et al. (35) elaborated significant relevant arguments against the BSA adjustment of GFR, most importantly that BSA indexing fails to equalize the GFR values of two subjects with different BSA. In a review article, Hsu et al. (36) concluded that there is no consensus as to whether or not to normalize GFR for BSA. Indexing the CG formula values for BSA is often referred to as eGFR, but this practice is questionable because weight is already included as a variable. Furthermore, Cockcroft and Gault stated that for patients with marked obesity or ascites, the absolute weight should be substituted with lean or ideal body weight (4). Recently, in a very large data set, lean body mass improved the prediction of eGFR (37,38). Macdonald et al. (39,40) demonstrated that muscle mass, compared with the MDRD equation variables, explained more variance in the mGFR and improved its estimation. If the change of weight is adding fat, correction for BSA lowers the GFR value artificially. Obesity is associated with hyperglycemia, microalbuminuria, and hypertension, all of which may initially cause glomerular hyperfiltration. After excluding subjects with these confounding factors, Anastasio et al. (41) observed no significant difference in absolute GFR between obese (body mass index [BMI] = 46.8 ± 1.78 kg/m², BSA = 2.18 ± 0.04 m²) and nonobese (BMI = 25.6 ± 0.93 kg/m², BSA = 1.76 ± 0.04 m²) subjects, but the BSA-adjusted GFR significantly decreased from 106.6 ± 3.0 ml/min to 84.1 ± 2.32 ml/min per 1.73 m² in the obese group. These findings have significant implications in view of the estimated 26.7% prevalence of obesity in the United States (42).

In clinical practice, for purposes of drug dosage, the absolute GFR should be used instead of the BSA-adjusted one, and this is recommended by the National Kidney Foundation (43). In the kidney transplant evaluation process, indexing for BSA may eliminate potential donors by artificially lowering their GFR. Conversely, adjusting the GFR of a small-sized donor will spuriously increase its actual value, but this would be irrelevant for a normal-sized recipient. Therefore, in these case scenarios, the BSA indexing of GFR is not advisable.

The MDRD and CKD-EPI equations are BSA-adjusted by intrinsic design, therefore “eliminating” the need for BSA indexing of their eGFR values. There are no data for mean BSA in the development group for CG formula. The MDRD and CKD-EPI equations included populations with mean BSA of 1.91 ± 0.23 and 1.93 ± 0.20 m² and BMI of 28 (calculated) and 28 ± 6 kg/m², respectively, revealing overall overweight populations (5,11). Reversing the BSA indexing process shows mean mGFR increase from 39.8 ml/min per 1.73 m² to 43.9 ml/min and 68.0 ml/min per 1.73 m² to 75.9 ml/min for the MDRD and CKD-EPI equation populations, respectively. Therefore the “BSA-adjusted” MDRD and CKD-EPI equations decreased the absolute mGFR of their included populations on average by −4.1 ml/min (−9.3%) and −7.9 ml/min (−10.4%), respectively. The BSA bias could be a reason for the need to refit the MDRD equation for Chinese and Japanese populations in view of their lower mean BSA of 1.7 ± 0.18 and 1.65 ± 0.19 m² and BMI of 23.6 ± 3.6 and 23 kg/m² (calculated), respectively (44,45). Therefore, large population samples with diverse body size need to be studied to determine what the best method for GFR indexing might be.

Figure 2. Difference (bias) between simultaneously measured GFRs by urinary clearances of intravenous infusion of inulin (Cin) and subcutaneous injection or intravenous infusion of 125I-iothalamate (Cio) at different Cin levels. Total bias by Bland-Altman plot (A), and relative bias at different Cin levels (B). The dashed lines represent mean difference and respective +2 and −2 SD limits. The data were derived from five studies including 94 subjects (23,27–30).
Normal GFR According to Age, Gender, and Race

Normal mGFR values in whites are approximately 130 ml/min per 1.73 m² for young men and 120 ml/min/1.73 m² for young women and decline as persons age (16,17). Renal senescence was reviewed in details by Glasscock et al. (46). Rule et al. (47), in an elegant study of kidney donor biopsies, revealed that nephrosclerosis increased with age, but it was not associated with CKD risk factors after adjustment for age and gender. Surprisingly, the observed mGFR decline with normal aging was not explained by nephrosclerosis. The K/DOQI-CKD defined GFR <60 ml/min per 1.73 m² as a disease, regardless of age, gender, or race. However, it was stated: “the cut-off levels between stages are inherently arbitrary.” Extrapolating individual data for 70 men without CKD, age 24 to 89 years (48), K/DOQI-CKD reported normal mean mGFR for age groups by decades on the basis of linear regression analysis, assuming that the values for women were 8% lower at all ages (2). For men age groups >50 and women age groups >40 years, the two-SD range is below 60 ml/min per 1.73 m²; therefore the K/DOQI-CKD classification will misdiagnose some of these subjects with CKD stage 3.

After Davies and Shock’s study (48), published 60 years ago, there is still limited data for Cin in normal subjects by age groups (49), particularly for the elderly (age >75 years). The significant limitations of this study are the inclusion of men only and the small number of elderly participants: nine in the age group 70 to 79 years and 12 in the age group 80 to 89 years. A large study of a healthy Chinese population with mGFR by ⁹⁹mTc-DTPA included a slightly higher number of elderly subjects: 49 in the age group of 70 to 80 years and 12 in the age group of >80 years (50). Three other large studies with Cio reported mGFR by age groups in kidney donors (presumably normal individuals) from the United States, but all of the individuals were aged <75 years (51–53), and only one study reported data for African Americans (53).

Obviously, large studies to establish the boundaries of normal mGFR according to a wide range of ages in different genders and races are needed to resolve recent debates (54–57) about overdiagnosing CKD in the elderly, a population of 16.6 million of >75 years according to the 2000 United States population census, and estimated to be 18.8 million as of 2009 (58). Hopefully, the ongoing Berlin Initiative Study in the elderly would provide some answers for proper CKD diagnosing in this population (59).

Reporting the eGFR

The K/DOQI-CKD panel should be commended for introducing the eGFR reporting by clinical laboratories. This has significant clinical implications because it reminded the practitioners that the SCr concentration becomes “abnormal” after up to 50% mGFR loss (14). Unfortunately, by requesting automatic laboratory eGFR reporting whenever SCr is ordered (60), the current K/DOQI-CKD and National Kidney Disease Education Program guidelines created several controversies in clinical practice that were extensively debated elsewhere (54–57,61–64). In 2007, the Australasian Creatinine Consensus Working Group revised their recommendation to extend the upper reporting limit of eGFR from 60 to 90 ml/min per 1.73 m² and appropriately advised the practitioners that eGFR 45 to 59 ml/min per 1.73 m² in people ≥70 years of age may not be associated with CKD, providing the eGFR was stable over time and no other CKD markers, such as proteinuria, were present (65). For mGFR >60 ml/min per 1.73 m², the new CKD-EPI equation, compared with the MDRD one, had a significantly lower median bias (~3.5 versus ~10.6 ml/min per 1.73 m², respectively) in a large validation data set. If this result is confirmed in other large populations, it will be helpful for the practitioner to change the current practice by many laboratories to report simply “eGFR >60 ml/min per 1.73 m².” For example, for a 30-year-old white woman with SCr of 1.1 mg/dl, the eGFR (CKD-EPI) is 68 ml/min per 1.73 m², representing approximately 40% loss of the normal mGFR. Even after accounting for the formulas’ underestimations at this level of mGFR, this is a significant eGFR loss, something of which the clinician would not be aware by looking at the eGFR >60 ml/min per 1.73 m² report. All combined, these new recommendations would prevent missing younger subjects with decreased eGFR in the range of 60 to 90 ml/min per 1.73 m² who might have early asymptomatic CKD and minimize inappropriate CKD diagnosing in the elderly. Furthermore, a recent cost-benefit analysis of the automatic eGFR reporting suggested “that reporting eGFR may be beneficial, but this limited benefit was reversed with virtually any reduction in quality of life caused by incorrect diagnosis of CKD” (66). Therefore, the eGFR report should be requested only by practitioners who are familiar with the pitfalls of its interpretation.

SCr Calibration: The Need for International Normalized Ratio for SCr Measurement

The National Kidney Disease Education Program Laboratory Working Group launched a global SCr standardization effort to reduce interlaboratory variation. Creatinine assays using the Jaffe reaction may have higher SCr values because of interference from noncreatinine chromogens, thus creating mostly underestimation of the high levels of eGFR (67). One study (68) revealed that the mean eGFR bias of the simplified (four variables) original MDRD equation (69) with Beckman CX3 kinetic Jaffe assay was reduced from 6.7% to 0.4% for the isotope dilution mass spectrometry (IDMS) assay, 8.8% to 3.4% for the Roche enzymatic assay, and 13.7% to 7.1% for the Roche compensated kinetic Jaffe assay when applying the re-expressed IDMS-traceable MDRD formula (70). Figure 3 shows the eGFR bias, range 0% to 24.2%, between the original and IDMS-traceable MDRD equations at different IDMS-traceable SCr levels. Interestingly, this eGFR calibration bias range is much lower than the Cio-Cin one (~37.1% to 61.3%), thus reinforcing the need of further research to “calibrate” any differences between various GFR markers.

The National Institute for Standards and Technology in the United States released only two levels of Standard Reference Material for SCr (SRM 967), 0.75 and 3.92 mg/dl, for calibrating the SCr measurements to a uniform standard, the IDMS. The actual process of standardization is at the discretion of each laboratory. Using only two point references to calibrate the entire spectrum of SCr values presumes that all calculated “calibrated” values would not
deviate from the identity line. An analytic CV of SCr values among laboratories as low as 2% translates into a critical difference of 12.4% and has significant clinical implications for estimating the GFR, especially in the low range of SCr levels (71). Furthermore, when three previously reported linear equations (72–74) and a proposed correction value (75) for SCr calibration to the MDRD Study laboratory were tested in 10,108 subjects, those methods of calibration did not show valid reproducible eGFR values for purposes of K/DOQI-CKD classification (76). Extensive calibration methods of SCr assays from several laboratories had very minimal or no effect on the MDRD equation’s accuracy in six of the nine studied populations (excluding the MDRD one) (77). Therefore, there is a need to establish an international normalized ratio for SCr measurement by including reference IDMS values for a broad SCr spectrum to improve the current standardization process.

**Evaluating the Performance of eGFR Formulas**

Evaluating the performance of prediction equations by mainly emphasizing the bias could be misleading, because a bias of zero does not assure perfect performance if the positive and negative eGFR differences are equally spread around the identity line of mGFR. Precision is a very helpful additional test to clinically interpret the bias. Total (ml/min per 1.73 m²) and relative (%) values for bias and precision have different implications at different GFR levels, because they should be interpreted in a particular clinical context. For example, small eGFR total bias and precision suggest a good performance; however, this could simply reflect low mGFR values, and high relative bias and precision may unmask poor performance. Accuracy, expressed as a percentage of eGFRs within a percentage range of mGFR, combines precision and bias by expressing how many GFR estimates are dispersed, i.e. the density, within a given range of their respective GFR measurements. Therefore, it is the most useful statistical analysis for the clinician. Because day-to-day mGFR variation was reported to be as high as 17% (3,23,78), the accuracy within ± 15% of mGFR provides the clinician with the percentage of eGFRs not deviating from the mGFRs.

PubMed database research (last accessed July 2010) for data sets reporting accuracy for the CG and MDRD formulas in adults with a minimum of 100 mGFRs, because estimates from small studies can be unreliable (2), recovered 61 publications. Studies with subsets of elsewhere-reported data sets were excluded. Increasing the minimal number of mGFRs to 300 narrowed the field to 25 publications (33 data sets), which are presented in Table 1(6,11,37,38,44,45,72,77,79–95). The most frequently reported accuracy range was within ± 30%, followed by within ± 50%, which has very little clinical application. Less than half of the studies provided accuracy within ± 15%, median bias, and precision by SD or interquartile range; therefore no analysis was performed for these results. The range of mean bias (eGFR–mGFR, calculated whenever not provided) was −9.0 to 13.6 and −29.2 to 12 ml/min per 1.73 m², and accuracy within ± 30% ranged from 49% to 88% and 53% to 89%, for the CG and MDRD equations, respectively. The number of mGFRs in the data sets varied from 322 to 5504. Thus, for comparison of both equations’ performance, after excluding three data sets with significant difference in the number of CG and MDRD estimates (80), a weighted average bias and accuracy were calculated as follows,

\[ \bar{x} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \]

where w is the number of mGFRs, and x is the bias or accuracy of eGFRs in a particular data set.

The weighted average bias was 0.02 versus −5.34 ml/min per 1.73 m² (6,45,72,80–87,89,92,93,95), and accuracy within ± 30% was 72% versus 78% (6,37,38,45,72,77,79–87,89,91–93,95), for the CG and MDRD formulas, respectively. The weighted average

![Figure 3](image-url)
Table 1. Performance of CG and MDRD formulas in studies with minimum of 300 measured GFRs

<table>
<thead>
<tr>
<th>Study</th>
<th>GFR Method</th>
<th>SCr Calibration</th>
<th>Population</th>
<th>( n )</th>
<th>mGFRs</th>
<th>Mean mGFR ± SD (Range)</th>
<th>50%</th>
<th>30%</th>
<th>15%</th>
<th>Mean</th>
<th>Median</th>
<th>50%</th>
<th>30%</th>
<th>15%</th>
<th>Mean</th>
<th>Median</th>
<th>Precision SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stevens et al. (77)</td>
<td>( ^{125} )I-iothalamate</td>
<td>Yes (IDMS)</td>
<td>Mixed</td>
<td>5504</td>
<td>68 ± 40</td>
<td>69 ± 83</td>
<td>5.5</td>
<td>6.0</td>
<td>6.0</td>
<td>6.0</td>
<td>6.0</td>
<td>5.5</td>
<td>6.0</td>
<td>6.0</td>
<td>6.0</td>
<td>6.0</td>
<td>5.5</td>
</tr>
<tr>
<td>Levey et al. (11)</td>
<td>( ^{125} )I-iothalamate</td>
<td>Yes (IDMS)</td>
<td>Mixed</td>
<td>3896</td>
<td>68 ± 36</td>
<td>81 ± 81</td>
<td>81</td>
<td>81</td>
<td>81</td>
<td>81</td>
<td>81</td>
<td>81</td>
<td>81</td>
<td>81</td>
<td>81</td>
<td>81</td>
<td>81</td>
</tr>
<tr>
<td>Botev et al. (6)</td>
<td>Inulin</td>
<td>No</td>
<td>Mixed</td>
<td>2208</td>
<td>72 ± 39</td>
<td>89 ± 94, 71 ± 69, 38 ± 34</td>
<td>-3.5</td>
<td>-3.5</td>
<td>-3.5</td>
<td>-3.5</td>
<td>-3.5</td>
<td>21.5</td>
<td>20.0</td>
<td>20.0</td>
<td>20.0</td>
<td>20.0</td>
<td>20.0</td>
</tr>
<tr>
<td>Froissart et al. (72)</td>
<td>( ^{51} )Cr-EDTA</td>
<td>No</td>
<td>Mixed</td>
<td>2095</td>
<td>61 ± 27</td>
<td>91 ± 96, 79 ± 87, 49 ± 58</td>
<td>1.9</td>
<td>1.9</td>
<td>1.9</td>
<td>1.9</td>
<td>1.9</td>
<td>15.4</td>
<td>13.7</td>
<td>13.7</td>
<td>13.7</td>
<td>13.7</td>
<td>13.7</td>
</tr>
<tr>
<td>K/DOQI-CKD (2) for Lewis et al. (79)</td>
<td>( ^{125} )I-iothalamate</td>
<td>Yes (ClClin)</td>
<td>CKD and KT donors</td>
<td>1775</td>
<td>57 ± 33</td>
<td>94 ± 98, 80 ± 88</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
</tr>
<tr>
<td>Gonwa et al. (80)</td>
<td>( ^{125} )I-iothalamate</td>
<td>Yes (ClClin)</td>
<td>Mixed</td>
<td>1447</td>
<td>90 ± 32</td>
<td>84 ± 86, 61 ± 67</td>
<td>-8.5</td>
<td>-8.5</td>
<td>-8.5</td>
<td>-8.5</td>
<td>-8.5</td>
<td>-8.5</td>
<td>-8.5</td>
<td>-8.5</td>
<td>-8.5</td>
<td>-8.5</td>
<td>-8.5</td>
</tr>
<tr>
<td>Ibrahim et al. (81)</td>
<td>( ^{125} )I-iothalamate</td>
<td>Yes (ClClin)</td>
<td>Mixed</td>
<td>1286</td>
<td>122 ± 23</td>
<td>97 ± 98, 88 ± 78</td>
<td>-6.0</td>
<td>-6.0</td>
<td>-6.0</td>
<td>-6.0</td>
<td>-6.0</td>
<td>-6.0</td>
<td>-6.0</td>
<td>-6.0</td>
<td>-6.0</td>
<td>-6.0</td>
<td>-6.0</td>
</tr>
<tr>
<td>Nyman et al. (37) and Bjork et al. (38)</td>
<td>( ^{125} )I-iothalamate</td>
<td>Yes (ClClin)</td>
<td>Mixed</td>
<td>850</td>
<td>55 (9 to 121)</td>
<td>79 ± 80</td>
<td>3.4</td>
<td>3.4</td>
<td>3.4</td>
<td>3.4</td>
<td>3.4</td>
<td>3.4</td>
<td>3.4</td>
<td>3.4</td>
<td>3.4</td>
<td>3.4</td>
<td>3.4</td>
</tr>
<tr>
<td>Verhave et al. (82)</td>
<td>( ^{99m} )Tc-DTPA</td>
<td>No</td>
<td>SCR &lt; 1.5 mg/dl</td>
<td>850</td>
<td>99 ± 20</td>
<td>98 ± 99, 87 ± 89</td>
<td>-4.9</td>
<td>-4.9</td>
<td>-4.9</td>
<td>-4.9</td>
<td>-4.9</td>
<td>-4.9</td>
<td>-4.9</td>
<td>-4.9</td>
<td>-4.9</td>
<td>-4.9</td>
<td>-4.9</td>
</tr>
<tr>
<td>Poggio et al. (83)</td>
<td>( ^{125} )I-iothalamate</td>
<td>Yes (ClClin)</td>
<td>Mixed</td>
<td>828</td>
<td>32 ± 28</td>
<td>77 ± 89, 60 ± 71</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
</tr>
<tr>
<td>Bosma et al. (84)</td>
<td>( ^{99m} )Tc-DTPA plasma</td>
<td>Yes (ClClin)</td>
<td>Mixed</td>
<td>457</td>
<td>106 ± 18</td>
<td>97 ± 98, 85 ± 86</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Ma et al. (44)</td>
<td>( ^{99m} )Tc-DTPA plasma</td>
<td>No</td>
<td>Mixed</td>
<td>798</td>
<td>55 ± 18</td>
<td>76 ± 88</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
</tr>
<tr>
<td>Gera et al. (85)a</td>
<td>( ^{125} )I-iothalamate</td>
<td>No</td>
<td>Mixed</td>
<td>684</td>
<td>55 ± 35</td>
<td>94 ± 66, 30 ± 78</td>
<td>-7.8</td>
<td>-7.8</td>
<td>-7.8</td>
<td>-7.8</td>
<td>-7.8</td>
<td>-7.8</td>
<td>-7.8</td>
<td>-7.8</td>
<td>-7.8</td>
<td>-7.8</td>
<td>-7.8</td>
</tr>
<tr>
<td>Saleem et al. (86)</td>
<td>( ^{99m} )Tc-DTPA plasma</td>
<td>No</td>
<td>Mixed</td>
<td>601</td>
<td>83 (4 to 214)</td>
<td>92 ± 73, 75 ± 17</td>
<td>-5.1</td>
<td>-5.1</td>
<td>-5.1</td>
<td>-5.1</td>
<td>-5.1</td>
<td>-5.1</td>
<td>-5.1</td>
<td>-5.1</td>
<td>-5.1</td>
<td>-5.1</td>
<td>-5.1</td>
</tr>
<tr>
<td>Rule et al. (88)a</td>
<td>( ^{99m} )Tc-DTPA plasma</td>
<td>No</td>
<td>Mixed</td>
<td>530</td>
<td>48 ± 25</td>
<td>75 ± 75</td>
<td>-5.5</td>
<td>-5.5</td>
<td>-5.5</td>
<td>-5.5</td>
<td>-5.5</td>
<td>-5.5</td>
<td>-5.5</td>
<td>-5.5</td>
<td>-5.5</td>
<td>-5.5</td>
<td>-5.5</td>
</tr>
<tr>
<td>Hejs et al. (89)</td>
<td>( ^{51} )Cr-EDTA plasma</td>
<td>No</td>
<td>Mixed</td>
<td>592</td>
<td>47 ± 34</td>
<td>63 ± 53</td>
<td>-4.9</td>
<td>-4.9</td>
<td>-4.9</td>
<td>-4.9</td>
<td>-4.9</td>
<td>25.8</td>
<td>17.2</td>
<td>17.2</td>
<td>17.2</td>
<td>17.2</td>
<td>17.2</td>
</tr>
<tr>
<td>Ma et al. (90)b</td>
<td>( ^{99m} )Tc-DTPA plasma</td>
<td>NA</td>
<td>Mixed</td>
<td>567</td>
<td>57.9 ± 36.4 91 ± 64, 43 ± 76</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td></td>
</tr>
<tr>
<td>Burke et al. (91)</td>
<td>( ^{99m} )Tc-DTPA plasma</td>
<td>NA</td>
<td>Mixed</td>
<td>546</td>
<td>47 ± 16</td>
<td>85 ± 91, 59 ± 73</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
</tr>
<tr>
<td>Mariat et al. (92)</td>
<td>( ^{125} )I-iothalamate</td>
<td>No</td>
<td>Mixed</td>
<td>476</td>
<td>47 ± 16</td>
<td>85 ± 91, 59 ± 73</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
<td>6.6</td>
</tr>
</tbody>
</table>
### Table 1. (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>GFR Method</th>
<th>SCr Calibration</th>
<th>Population</th>
<th>n</th>
<th>mGFRs</th>
<th>Mean mGFR ± SD (Range)</th>
<th>Accuracy</th>
<th>Bias</th>
<th>Precision SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issa et al. (93)</td>
<td>$^{125}$I-iothalamate</td>
<td>Yes (IDMS)</td>
<td>KT donors</td>
<td>423</td>
<td>106</td>
<td>88 ± 16</td>
<td>50%</td>
<td>2.0</td>
<td>−16.0 −0.5 −16.3</td>
</tr>
<tr>
<td>Matsuo et al. (45)$^c$</td>
<td>Inulin</td>
<td>Yes (CKClin)</td>
<td>CKD, KT donors</td>
<td>413</td>
<td>59.1</td>
<td>76 ± 35.4</td>
<td>30%</td>
<td>7.0</td>
<td>12.0 −19.6 22.2</td>
</tr>
<tr>
<td>Horio et al. (94)$^d$</td>
<td>Inulin</td>
<td>Yes (IDMS)</td>
<td>CKD</td>
<td>390</td>
<td>45.2</td>
<td>75 ± 34.7</td>
<td>15%</td>
<td>−1.3</td>
<td>19.4</td>
</tr>
<tr>
<td>Tidman et al. (95)</td>
<td>Iohexol plasma</td>
<td>Yes (IDMS)</td>
<td>NA</td>
<td>322</td>
<td>50.4</td>
<td>85 ± 28.1</td>
<td>Mean</td>
<td>5.9</td>
<td>−1.0 13.2 12.9</td>
</tr>
</tbody>
</table>

The data in bold type were included for average weighted analysis (see text). CG, Cockcroft-Gault formula; MDRD, Simplified (4 variables) Modification of Diet in Renal Disease Study equation; mGFR, measured GFR in ml/min per 1.73m², urinary clearance unless otherwise described; $^{51}$Cr-EDTA, $^{99m}$Tc-DTPA, $^{99m}$Technetium-diethyleneetriamine penta-acetate; SCr, serum creatinine, calibrated to ClClin (Cleveland Clinic Laboratory) or IDMS (isotope dilution mass spectroscopy); CKD, chronic kidney disease; KT, kidney transplant; DM, diabetes mellitus; mGFR, measured GFR in ml/min per 1.73 m²; Accuracy, percentage of GFR estimates within ±50%, ±30%, and ±15% of mGFRs; Bias, estimated minus measured GFR in ml/min per 1.73 m² (mean bias calculated if possible whenever not provided); Precision, one standard deviation (SD) of mean bias in ml/min per 1.73 m²; estimated GFR by CG in ml/min per 1.73 m² (ml/min in references 37,80,84,85), data in italic font for corrected CG by adjusted (37) or lean body weight (86), or 0.84 for creatinine tubular secretion (87); estimated GFR by MDRD in ml/min per 1.73 m², data in italic for IDMS re-expressed MDRD equation.

$^a$Study design included SCr measurements not exactly on the day of GFR measurement.
$^b$MDRD modified for Chinese.
$^c$CG modified for Japanese.
$^d$MDRD modified for Japanese.
accuracy within ±30% in data sets with mean mGFR >60 ml/min per 1.73 m² was 74% versus 80% and in data sets with mean mGFR <60 ml/min per 1.73 m² was 70% versus 75%, for the CG and MDRD formulas, respectively. The weighted average accuracy within ±30% in data sets with calibrated SCr was 74% versus 81%, and in data sets with noncalibrated SCr was 70% versus 74%, for the CG and MDRD formulas, respectively. Because the data sets had different markers for mGFR, a confounding factor for the MDRD equation, further analysis showed that the MDRD weighted accuracy within ±30% in studies with mGFR by Cio was 78%, similar to the performance of 76% in all 33 data sets.

The CKD-EPI equation, compared with the MDRD one, had 3.6% better overall accuracy in a large validation data set (11). Only six studies (two of those with >300 mGFRs) were recovered to report accuracy within ±30% for the new CKD-EPI equation, range 59% to 89% (11,94,96–99).

Unfortunately, there is no agreement in the literature on what constitutes excellent accuracy. Applying the widely accepted statistical analysis of 95% confidence interval (CI) as a criterion would imply that a formula has an excellent accuracy when 95% of its GFR estimates are within a ±15% range of their respective GFR measurements.

Recently, for graphic presentation of the mGFR and eGFR relation, some authors proposed plotting the mGFR

Figure 4. Different trend line (in bold) appearances when measured (by inulin) and estimated (by Cockcroft-Gault [CG] formula and Modification Diet in Renal Disease Study [MDRD] equation) GFR values are plotted on X and Y axes (A and B) and Y and X axes (C and D), respectively. Only the trend lines in A and B illustrate correctly the reported results for difference (bias) between estimated and measured GFR in the study (6). The MDRD equation was simplified to four variables. r², variability coefficient.
### Table 2. Classification performance of CG and MDRD formulas in mGFR groups according to K/DOQI-CKD Stages

<table>
<thead>
<tr>
<th>Study</th>
<th>mGFR Method</th>
<th>SCr Calibration</th>
<th>Population</th>
<th>Overall</th>
<th>mGFR Groups in ml/min per 1.73 m² according to K/DOQI-CKD Stages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>% eGFR</td>
<td>mGFR</td>
</tr>
<tr>
<td>CG formula</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Botev et al. (6)</td>
<td>Inulin</td>
<td>No</td>
<td>Mixed</td>
<td>62</td>
<td>1360</td>
</tr>
<tr>
<td>Froissart et al. (72)</td>
<td>51Cr-EDTA</td>
<td>Yes (ClClin)</td>
<td>CKD and KT donors</td>
<td>68</td>
<td>1416</td>
</tr>
<tr>
<td>Mariat et al. (92)</td>
<td>Inulin</td>
<td>No</td>
<td>KT recipients</td>
<td>57</td>
<td>270</td>
</tr>
<tr>
<td>Tidman et al. (95)</td>
<td>Iohexol plasma</td>
<td>Yes (IDMS)</td>
<td>N/A</td>
<td>64</td>
<td>206</td>
</tr>
<tr>
<td>Chudleigh et al. (103)</td>
<td>51Cr-EDTA</td>
<td>No</td>
<td>DM type 2</td>
<td>70</td>
<td>202</td>
</tr>
<tr>
<td>Ma et al. (99)*</td>
<td>Ilo-thalamate</td>
<td>Yes (IDMS)</td>
<td>KT donors and others</td>
<td>63</td>
<td>170</td>
</tr>
<tr>
<td>Rostoker et al. (104)</td>
<td>Inulin</td>
<td>No</td>
<td>CKD</td>
<td>72</td>
<td>193</td>
</tr>
<tr>
<td>Rigalkeau et al. (105)</td>
<td>51Cr-EDTA</td>
<td>No</td>
<td>DM</td>
<td>55</td>
<td>110</td>
</tr>
<tr>
<td>White et al. (106)</td>
<td>99mTc-DTPA</td>
<td>Yes (CI Clin)</td>
<td>N/A</td>
<td>90</td>
<td>136</td>
</tr>
<tr>
<td>Urbaniak et al. (107)</td>
<td>Iohexol plasma</td>
<td>No</td>
<td>Mixed</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>All studies</td>
<td></td>
<td></td>
<td></td>
<td>64</td>
<td>4129</td>
</tr>
<tr>
<td>MDRD equation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Botev et al. (6)</td>
<td>Inulin</td>
<td>No</td>
<td>Mixed</td>
<td>62</td>
<td>1261</td>
</tr>
<tr>
<td>Froissart et al. (72)</td>
<td>51Cr-EDTA</td>
<td>Yes (ClClin)</td>
<td>CKD and KT donors</td>
<td>71</td>
<td>1483</td>
</tr>
<tr>
<td>Ma et al. (90)*</td>
<td>99mTc-DTPA</td>
<td>Yes (ClClin)</td>
<td>N/A</td>
<td>65</td>
<td>370</td>
</tr>
<tr>
<td>Mariat et al. (92)</td>
<td>Inulin</td>
<td>No</td>
<td>KT recipients</td>
<td>63</td>
<td>298</td>
</tr>
<tr>
<td>Tidman et al. (95)</td>
<td>Iohexol plasma</td>
<td>Yes (IDMS)</td>
<td>N/A</td>
<td>62</td>
<td>200</td>
</tr>
<tr>
<td>Chudleigh et al. (103)</td>
<td>51Cr-EDTA</td>
<td>No</td>
<td>DM type 2</td>
<td>63</td>
<td>202</td>
</tr>
<tr>
<td>Ma et al. (44)</td>
<td>99mTc-DTPA</td>
<td>Yes (ClClin)</td>
<td>CKD</td>
<td>29</td>
<td>84</td>
</tr>
<tr>
<td>Michels et al. (99)*</td>
<td>Ilo-thalamate</td>
<td>Yes (IDMS)</td>
<td>KT donors and others</td>
<td>65</td>
<td>176</td>
</tr>
</tbody>
</table>
Table 2. (Continued)

<table>
<thead>
<tr>
<th>mGFR Groups in ml/min per 1.73 m² according to K/DOQI-CKD Stages</th>
<th>Study</th>
<th>SCr Method</th>
<th>CG Formula</th>
<th>MDRD Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;0.15</td>
<td>90</td>
<td>% eGFR mGFR</td>
<td>% eGFR mGFR</td>
<td>% eGFR mGFR</td>
</tr>
<tr>
<td>15–29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30–59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60–69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;90</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The data for the percentage or number of eGFRs were extrapolated if not provided. CG, Cockcroft-Gault formula; MDRD, Simplified (4 variables) Modification of Diet in Renal Disease Study equation; mGFR, measured GFR in ml/min per 1.73 m², urinary clearance unless otherwise described; K/DOQI-CKD, Kidney Disease Outcome Quality Initiative Clinical Practice Guidelines for Chronic Kidney Disease; eGFR, estimated GFR in ml/min per 1.73 m² by CG and MDRD. SCr calibration and 62% and 59% without SCr calibration, for eGFR assignment was 67% and 65% in studies with SCr calibration and 62% and 59% without SCr calibration, for the CG and MDRD formulas, respectively. Only one study was found to evaluate the CKD-EPI equation in GFR groups by mGFR, and its proper eGFR overall assignment was 69% (99). These results suggest that eGFR prediction equations should be used with great caution for CKD diagnosing or staging given the clinical and therapeutic consequences for millions of people, i.e. approximately 10 million (38%) subjects might have been misclassified by the MDRD equation on the basis of estimated CKD prevalence (4,5,11), on the basis of statistically significant regression coefficients for those variables in relation to the mGFR. These models may not account for variations in the tubular handling of creatinine (15). Hence, mGFR is the independent variable, and eGFR is the dependent one. Therefore, mGFR values should be plotted on the X axis, and eGFR values should be plotted on the Y axis, as illustrated in a similar model (102).

Effect of eGFR on K/DOQI-CKD Diagnosing and Staging

For the practicing clinician, the ultimate evaluation of any formula’s performance is its proper eGFR assignment into the five GFR groups of the K/DOQI-CKD classification where the GFR groups are defined by mGFR. PubMed research (last accessed July 2010) recovered only 12 studies with mGFRs >10 reporting this performance of the CG and MDRD formulas, and their results are presented in Table 2 (6,44,72,90,92,95,99,103–107). Comparing CG versus MDRD formula in studies with data for both formulas (6431 subjects) showed similar overall proper classification of the subjects, 64% versus 63%, range 55% to 72% versus 53% to 71%, respectively (6,72,92,95,99,103–107). The best performance of CG and MDRD equations was in mGFR group 30 to 59 ml/min per 1.73 m², whereas the worst one was in mGFR groups <15 and >90 ml/min per 1.73 m², respectively. Proper overall eGFR assignment was 67% and 65% in studies with SCr calibration and 62% and 59% without SCr calibration, for the CG and MDRD formulas, respectively. Only one study was found to evaluate the CKD-EPI equation in GFR groups by mGFR, and its proper eGFR overall assignment was 69% (99). Some authors proposed that the CKD staging should be done by eGFR, as a category of evaluation, instead of mGFR for two main reasons: mGFR is not identical to true GFR because of errors in the measurement process, and the objective in development of eGFR equations is to determine an eGFR that is unbiased for the mGFR for populations with a given level of eGFR (101). First, although the GFR measurement by calculated clearance of a given substance may vary from the actual GFR, it is the best way to assess renal function.
because of errors in the measurement process, there are currently no other practical means to measure the glomerular filtration rate. Similarly, the GFR estimation is also affected by errors in the SCr measurement because if the same sample were measured repeatedly, a CV range from 1% to 11% has been reported (36). Any repeated measurements have some errors related to the process of measurement, and the solution is to minimize the errors by standardizing the process instead of substituting the measurement with an estimate, especially when the GFR deviates >30% from the mGFR in approximately 10% to 50% of the estimates. Second, the purpose of calculating eGFR is to avoid the cumbersome measurement of GFR in daily practice. For diagnosis and classification of CKD, the clinician is primarily interested in the reliability of estimating the mGFR. Therefore, CKD staging (GFR stratification) by GFR and then evaluating the performance of an eGFR equation in the thus-created stages is not clinically helpful, because it omits the mGFR confirmatory test.

Conclusion
The introduction of the K/DOQI-CKD classification was a landmark event allowing CKD to be recognized as a global public health problem. Estimating the GFR is now an integral part of the daily clinical practice, and it is used routinely for evaluation and monitoring of the renal function. The clinician should be aware of the significant limitations of the CG and MDRD formulas for purposes of diagnosing and staging CKD in the individual patient because of significant errors and significant misclassification.

Because current SCr-based formulas are derived from a particular population with a specific marker of mGFR, reference IDMS values for a broad SCr spectrum need to be established, and further research is required to define calibrating coefficients among different GFR markers and appropriate GFR indexing for subjects with different body sizes. This would improve the formulas’ performance in daily practice and allow eGFR comparisons between different populations in research studies. After calibration of the mGFR data from various studies of healthy subjects, samples of age-, gender-, and race-related normal ranges of mGFR should be established and incorporated in the laboratories’ eGFR report. Then what constitutes CKD by decreased levels of mGFR without kidney pathology markers (such as proteinuria) can be defined.

All combined, this should prevent underdiagnosing CKD by revealing early decreased GFR in young individuals, those who will benefit the most from an early nephrology referral, and avoid over-diagnosing CKD in elderly individuals by medicalization of renal senescence. Great caution should be exercised when embracing new definitions of disease in certain populations to preserve the first rule of practicing medicine: *primum non nocere*.

Disclosures
None.

References


62. Levey AS, Stevens LA, Hostetter T: Automatic reporting of


Published online ahead of print. Publication date available at www.cjasn.org.