A Comparison of Prediction Equations for Estimating Glomerular Filtration Rate in Adults without Kidney Disease

JULIE LIN, ERIC L. KNIGHT, MARY LOU HOGAN, and AJAY K. SINGH
Renal Division, Brigham and Women’s Hospital, Boston, Massachusetts

Abstract. The ability of the Modification of Renal Disease (MDRD) equation to predict GFR when compared with multiple other prediction equations in healthy subjects without known kidney disease was analyzed. Between May 1995 and December 2001, a total of 117 healthy individuals underwent 125I-iothalamate or 99mTc-diethylenetriamine-pentaacetic acid (DTPA) renal studies as part of a routine kidney donor evaluation at either Brigham and Women’s Hospital or Boston Children’s Hospital. On chart review, 100 individuals had sufficient data for analysis. The MDRD 1, MDRD 2 (simplified MDRD equation), Cockcroft-Gault (CG), Cockcroft-Gault corrected for GFR (CG-GFR), and other equations were tested. The median absolute difference in ml/min per 1.73 m² between calculated and measured GFR was 28.7 for MDRD 1, 18.5 for MDRD 2, 33.1 for CG, and 28.6 for CG-GFR in the 125I-iothalamate group and was 31.1 for MDRD 1, 38.2 for MDRD 2, 22.0 for CG, and 31.1 for CG-GFR in the 99mTc-DTPA group. Bias was −0.5, −3.3, 25.6, and 5.0 for MDRD 1, MDRD 2, CG, and CG-GFR, respectively, in subjects who received 125I-iothalamate and −33.2, −36.5, 6.0, and −15.0 for MDRD 1, MDRD 2, CG, and CG-GFR, respectively, in those who received 99mTc-DTPA studies. Precision testing, as measured by linear regression, yielded R² values of 0.04 for CG, 0.05 for CG-GFR, 0.15 for MDRD 1, and 0.14 for MDRD in those who underwent 125I-iothalamate studies and 0.18 for CG, 0.21 for CG-GFR, 0.40 for MDRD 1, and 0.38 for MDRD 2 for those who underwent 99mTc-DTPA studies. The MDRD equations were more accurate within 30 and 50% of the measured GFR compared with the CG and CG-GFR equations. When compared with the CG equation, the MDRD equations are more precise and more accurate for predicting GFR in healthy adults. The MDRD equations, however, consistently underestimate GFR, whereas the CG equations consistently overestimate measured GFR in people with normal renal function. In potential kidney donors, prediction equations may not be sufficient for estimating GFR; radioisotope studies may be needed for a better assessment of GFR. Further studies are needed to derive and assess GFR prediction equations in people with normal or mildly impaired renal function.

A noninvasive and accurate estimation of GFR is one of the holy grails of nephrology. Not only are prediction equations crucial for estimating GFR or creatinine clearance (CrCl) in the clinical research setting where only a single blood test is needed to derive and assess GFR prediction equations in people with normal or mildly impaired renal function. In potential kidney donors, prediction equations may not be sufficient for estimating GFR; radioisotope studies may be needed for a better assessment of GFR. Further studies are needed to derive and assess GFR prediction equations in people with normal or mildly impaired renal function.

Materials and Methods

Between May 1995 and December 2001, a total of 117 healthy adults underwent 125I-iothalamate or 99mTc-diethylenetriamine-pentaacetic acid (DTPA) renal clearance studies as part of a routine work-up for potential kidney donation at either Brigham and Women’s Hospital (BWH) or Boston Children’s Hospital (Boston, MA). On chart review, a total of 100 subjects had sufficient clinical and laboratory data to estimate GFR, including age, gender, race, weight, height, and SCr. Those with missing data for blood urea nitrogen (BUN; n = 4) and serum albumin (n = 10) were given the default values of 15 mg/dl and 4.0 g/dl, respectively; the rationale is that these two variables together contribute <1% to the observed variance of the calculations (A.S. Levey, personal communication). Of the 100 subjects included in the study, 55 underwent 125I-iothalamate studies and 45 underwent 99mTc-DTPA.

Subjects who underwent 125I-iothalamate studies at BWH were asked to fast for at least 8 h and given a water load of 10 ml/kg and...
5 drops of potassium iodide diluted in 15 ml of water orally (to block thyroid uptake of 125I-iothalamate) at the initiation of the study. Thirty-five micro-Curies of 125I-iothalamate was injected subcutaneously into the upper arm. Blood was drawn and urine was sampled at time 0 (before 125I-iothalamate injection) and at 60, 120, and 180 min. Total urine volume and urinary flow rates were assessed every 60 min. Oral fluid hydration was administered at 500 ml/h as tolerated. GFR measurements for two timed urine collections were averaged and standardized for a body surface area (BSA) of 1.73 m². Intra-assay coefficients of variation (CV) were ≤10% for 40 of the 55 iothalamate subjects and ≥10% for the remaining 15 subjects. Forty-six subjects had CV of ≥15%. For those with CV ≥15%, the recorded minimum urinary flow rates ranged from 1.1 to 2.8 ml/min. The acceptable minimum urinary flow rate in the original protocol was 3 ml/min; if this was not achieved at time 60 min, then urine was collected at time 90 min and urinary flow rate was calculated at this time.

Individuals who underwent ⁹⁹ᵐTc-DTPA studies were hydrated orally or intravenously at 10 ml/kg per h for 30 min before study initiation. ⁹⁹ᵐTc-DTPA dosed at 50 μCi/kg was injected intravenously, and blood was sampled at 120, 180, and 240 min. Three GFR measurements were averaged and standardized for a BSA of 1.73 m². No information on intra-assay CV was available for ⁹⁹ᵐTc-DTPA studies.

Forty-one of 45 subjects who had ⁹⁹ᵐTc-DTPA studies had their SCr levels measured from their primary care provider’s office at a wide variety of laboratories; the remaining four patients had SCr assayed at the BWH laboratory. For subjects who underwent ¹²⁵I-iothalamate testing, 45 of 55 (82%) SCr levels were assayed at the BWH, and the remainder were obtained through different outside laboratories. The BWH laboratory used alkaline picrate reactions with a DAX96 (Bayer) machine to measure SCr through January 31, 2000, and an Olympus 640/2700 machine from February 1, 2000, to the present. The reference normal values for SCr were 0.8 to 1.6 mg/dl in men and 0.7 to 1.3 mg/dl in women present. The reference normal values for SCr were 0.8 to 1.6 mg/dl in men and an Olympus 640/2700 machine from February 1, 2000, to the present. The BWH laboratory used alkaline picrate reactions with a DAX96 (Bayer) machine to measure SCr through January 31, 2000, and an Olympus 640/2700 machine from February 1, 2000, to the present.

The prediction equations that we used are listed as follows:

1. Cockcroft-Gault (CG) (9): CrCl = (140 – Age) × Weight (kg)/SCr × 72
   a. For men: CrCl = [(140 – Age) × Weight (kg)/SCr × 72] × 0.85
2. CG-GFR estimate: GFR = 0.84 × CrCl by Equation (1)
3. MDRD 1 (2): GFR = 170 × [SCr]⁻⁰.⁹⁹⁹ × [Age]⁻⁰.₁⁷⁶ × [0.₇₆₂ if patient is female] × [1.₁₈ if patient is black] × [BUN]⁻⁰.₁₇₀ × [Alb]⁰.₃₁₈
4. MDRD 2 (8): GFR = 186 × [SCr]⁻¹.₄₅₄ × [Age]⁻⁰.₂₀₃ × [0.₇₄₂ if patient is female] × [1.₂¹₂ if patient is black]
5. Jelliffe 1 (X BSA/1.73 m²) (10)
   a. For men: (98 – [0.₈ × (age – 20)])/SCr
   b. For women: (98 – [0.₈ × (age – 20)])/SCr × 0.₉₀
6. Jelliffe 2 (11)
   a. For men: (100/SCr) – 12
   b. For women: (80/SCr) – 7
7. Mawer (12)
   a. For men: weight × [29.₃ – (0.₂₀₃ × age)] × [1 – (0.₀₃ × SCr)]
   (1.₄ × SCr) × (70/weight)
   b. For women: weight × [25.₃ – (0.₁₇₅ × age)] × [1 – (0.₀₃ × SCr)]

8. Bjornsson (13)
   a. For men: [27 – (0.₁₇₃ × age)] × weight × 0/SCr
   b. For women: [25 – (0.₁₇₅ × age)] × weight × 0.₀₇/SCr
9. Gates (14)
   a. For men: (89.₄ × SCr⁻¹.₂) + (55 – age) × (0.₄₄₇ × SCr⁻¹.₁)
   b. For women: (89.₄ × SCr⁻¹.₂) + (55 – age) × (0.₄₄₇ × SCr⁻¹.₁)
10. Salazar-Corcoran (15)
    a. For men: [137 – age] × [(0.₂₈₅ × weight) + (12.₁ × height²)]/ (51 × SCr)
    b. For women: [146 – age] × [(0.₂₈₇ × weight) + (9.₇₄ × height²)]/(60 × SCr)

Fisher’s exact test was used for proportions and t test for comparison of means. A Pearson correlation coefficient was also calculated for the normally distributed GFR data (Shapiro-Wilk test for normality, P = 0.₉₈).

The mean and median absolute differences were calculated from absolute difference = predicted value – measured value. The % absolute difference was calculated as % absolute difference = predicted value – measured value × 100 measured value.

Bias, a measure of systematic error, was defined by the mean prediction error (ME):

\[ N \]

\[ ME = \sum (pe_i)/N \]

\[ i = 1 \]

where pe_i = predicted value – true value and n = sample size.

The R² statistic was derived by simple linear regression (PROC GLM) and reflects the predictive ability of the model. P = 0.₅₅ was considered significant.

Refitting of the log-linear MDRD equation to our data set was performed with the following code using PROC REG in SAS:

For MDRD 1: Model ln(GFR) = ln(SCr) + ln(AGE) + ln(BUN) + ln(albumin) + race + gender

For MDRD 2: Model ln(GFR) = ln(SCr) + ln(AGE) + race + gender

SAS for Windows version 8.0 (Cary, NC) was used for all statistical calculations. Data collection by chart review was approved by the BWH Institutional Review Board.

Results

Clinical and laboratory characteristics of this study population are summarized in Table 1. The mean measured GFR was 112.₈ ml/min per 1.₇₃ m² (range, 70.₀ to 1₆₉.₀ ml/min). When data from subjects who received ¹²⁵I-iothalamate were compared with those who received ⁹⁹ᵐTc-DTPA, statistically significant differences were found in distribution of black race, mean weight, albumin, and measured GFR (Table 1).

Performance assessment for several CrCl and GFR equations are presented in Table 2. Mean absolute difference ranged from 2₆.₀ to ₄₉.₆, bias from -₁₈.₄ to ₂₈.₇, median absolute difference from 1₉.₀ to ₃₅.₄, median % absolute difference from 1₇ to ₃₀%, precision from ₀.₀₁ to ₀.₀₈, and Pearson correlations from ₀.₁₄ to ₀.₂₈.
values within 30% of the measured value ranged from 50 to 76%, whereas the percentage within 50% of measured values ranged from 72 to 96%. Results of CG compared with CG-GFR and MDRD 1 compared with MDRD 2 were very similar.

When the predicted values were compared with measured values stratified by the type of study (125 I-iothalamate or 99m Tc-DTPA), there was marked improvement in the MDRD predictions (Table 3). For those who received 125 I-iothalamate studies, the mean, median, and median % absolute differences did not change greatly, but the bias, precision, Pearson correlation, and accuracy improved with both MDRD equations. For those who received 99m Tc-DTPA studies, the mean, median, and median % absolute differences were increased in the MDRD 2 equation, whereas bias, precision, and Pearson correlation increased for both MDRD 1 and 2 equations and accuracy remained approximately the same (Table 3) when compared with the results for the combined 125 I-iothalamate and 99m Tc-DTPA studies (Table 2).

Because there were many overweight people who might have had a supraphysiologic calculated CrCl by CG, we also restricted the analyses to those with body mass index ≤30 (n = 73). No improvements were noted in any of the parameters examined in any of the equations (data not shown).

Refitting of the MDRD equation parameter estimates to our data set resulted in the following prediction equations:

MDRD 1: GFR = 278 × [SCr]−0.107 × [Age]−0.169 × [0.94 if patient is female] × [0.91 if patient is black] × [BUN]−0.089 × [Alb]−0.025 (R² = 0.11)

MDRD 2: GFR = 214 × [SCr]−0.113 × [Age]−0.174 × [0.96 if patient is female] × [0.92 if patient is black] (R² = 0.08)

Discussion

The MDRD and CG equations are the most widely recommended and used formulas for assessment of renal clearance; therefore, this discussion focuses primarily on these two equations. On the basis of a comparison of these findings to those in the published literature (Table 4), our first observation is that the MDRD equations perform much more poorly in subjects without kidney disease than in those with chronic kidney disease. This is not surprising since the MDRD equation was derived in MDRD study participants who were selected as having moderate to severe renal failure (measured 125 I-iothalamate GFR mean was 39.8 ml/min per 1.73 m²) (2).

Second, the MDRD prediction equations seem to systematically underestimate GFR, as indicated by the high negative bias (Table 3), especially as measured by the 99m Tc-DTPA method. Although the 99m Tc-DTPA group had a significantly higher mean weight, they also had significantly higher GFR measurements; therefore, these differences between the two groups are unlikely to explain the greater negative bias.

Third, we have confirmed that the simplified MDRD 2 equation loses very little predictive ability when compared with MDRD 1. Because serum albumin and BUN measurements may not be readily available, especially in the research setting, MDRD 2 is the prediction equation of choice over MDRD 1 in these situations. Likewise, estimating GFR from the CG CrCl calculation does not appreciably improve the predictive ability of the CG equation.

Last, 125 I-iothalamate and 99m Tc-DTPA techniques do not seem to be highly correlated with each other because the degree of bias, precision, linear correlation, and accuracy varied notably when the results were pooled (Table 2) rather than stratified by type of GFR testing (Table 3). The decreased precision observed when all 100 subjects were pooled for analysis may reflect the interassay variability that was conferred by using two different measurements of GFR for reference. However, high correlations of both 125 I-iothalamate and 99m Tc-DTPA to simultaneous measures of inulin clearance have been reported (r > 0.90) (16), although 125 I-iothalamate measurements exceeded inulin measurements by 14.6 to 25.9 ml/min per 1.73 m² (16) and 99m Tc-DTPA overestimated inulin clearance by 3.5 to 13.5 ml/min per 1.73 m² for (16,17). Therefore, analyses of prediction equations must take into account the type of GFR measurement, and caution should be used when comparing results of such investigations using diverse GFR measurement techniques. In this study, other

Table 1. Patient characteristics

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Combined</th>
<th>125 I-Iothalamate</th>
<th>99mTc-DTPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>100</td>
<td>55</td>
<td>45</td>
</tr>
<tr>
<td>Male gender</td>
<td>41 ± 10</td>
<td>21 (38%)</td>
<td>21 (47%)</td>
</tr>
<tr>
<td>Black race</td>
<td>10 (10%)</td>
<td>9 (16%)</td>
<td>1 (2%)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>78.1 ± 15.7</td>
<td>75.3 ± 15.7</td>
<td>81.5 ± 15.3</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>168.5 ± 9.4</td>
<td>168.8 ± 9.2</td>
<td>168.1 ± 9.8</td>
</tr>
<tr>
<td>Body surface area (m²)</td>
<td>1.9 ± 0.2 (1.4-2.4)</td>
<td>1.9 ± 0.2 (1.4-2.4)</td>
<td>1.9 ± 0.2 (1.4-2.3)</td>
</tr>
<tr>
<td>Serum creatinine (mg/dl)</td>
<td>0.9 ± 0.2 (0.5-1.3)</td>
<td>0.8 ± 0.2 (0.5-1.2)</td>
<td>0.9 ± 0.2 (0.5-1.3)</td>
</tr>
<tr>
<td>Serum urea nitrogen (mg/dl)</td>
<td>13 ± 4 (5-31)</td>
<td>13 ± 4 (6-31)</td>
<td>13 ± 4 (5-26)</td>
</tr>
<tr>
<td>Albumin (gm/dl)</td>
<td>4.5 ± 0.5 (3.3-5.4)</td>
<td>4.5 ± 0.4 (3.7-5.4)</td>
<td>4.3 ± 0.5 (3.3-5.4)</td>
</tr>
<tr>
<td>Measured GFR (ml/min per 1.73 m²)</td>
<td>112.8 ± 21.0 (70.0-169.0)</td>
<td>102.8 ± 15.8 (70.0-151.9)</td>
<td>125.1 ± 20.3 (77.3-169.0)</td>
</tr>
</tbody>
</table>

Results expressed as mean ± SD (range) or n (%).

p < 0.05 for iothalamate versus 99mTc-DTPA group.
factors that may have influenced the different correlations between $^{125}\text{I-}$iothalamate and $^{99m}\text{Tc-DTPA}$ include the higher
numbers of blacks and lower mean weights in the $^{125}\text{I-}$iothalamate group.

The MDRD equations had the least bias and highest accuracy when compared with $^{125}\text{I-}$iothalamate measurements. This is not
surprising considering that these equations were derived using $^{125}\text{I-}$iothalamate GFR measurements. The CG equation is un-
derstandably poorer, because it was derived by CrCl calculated by
24-h urinary creatinine collections as the “gold standard” in 249
adults ages 18 to 92 with mean SCr 0.99 to 1.78 mg/dl (9). In this
original study, 96% of study subjects were male, and no informa-
tion on race was given; this raises the issue of the generalizability
of the CG prediction equations. The use of 24-h urine collection
as the “gold standard” was also suboptimal because multiple
investigations have reported the inaccuracies of using 24-h urine
collections to measure CrCl, usually from under- or overcollection
by the subject (5,18,19). In fact, Coresh et al. (20) concluded that
there was no advantage in 24-h urine collection over the CG
estimations when compared with GFR determined by $^{125}\text{I-}$
iothalamate clearance.

Hsu et al. (21) reported finding systematic differences in SCr levels measured at BWH among those measured during
1997 and during 1998 when compared with values measured before 1997; therefore, the investigators adjusted their analysis
by adding 0.1 mg/dl to SCr values measured in 1997 and 0.3
mg/dl to values measured in 1998. In our data set, 35 of the 45
SCr assayed at BWH in the $^{125}\text{I-}$iothalamate group were per-
formed in 1998 or later, whereas all four of the SCr assayed at
BWH in the $^{99m}\text{Tc-DTPA}$ group were performed after 1998.
Because the vast majority of $^{125}\text{I-}$iothalamate patients were
assayed between 1998 and 2000 and because there is no
information (either published or provided by the BWH labo-
ry) of how SCr assays after 1998 are related to those
performed before 1998, we did not attempt to adjust SCr and
recalculate the prediction equations, although this is another
likely source of measurement error.

The issue of calibration of SCr laboratory measurement is

| Table 2. Mean calculated CrCl or GFR, mean absolute difference, bias, precision, and accuracy of GFR prediction equations compared with measured GFR$^a$ |
|---------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                     | Mean CrCl/GFR$^b$ (Range) | Mean Absolute Difference | Median Absolute Difference | Median % Absolute Difference | Bias$^c$ | $R^2$ (Precision) | Pearson Correlation | Accuracy $\%$ within 30% 50% |
| CG$^a$              | 129.6 ± 48.8 (50.1–288.5) | 37.9 | 26.7 | 23% | 16.8 | 0.06 | 0.24 | 58% 79% |
| CG-GFR$^a$          | 108.9 ± 41 (42.1–242.3) | 32.9 | 30.5 | 25% | −4.0 | 0.06 | 0.25 | 59% 83% |
| MDRD 1$^f$          | 97.6 ± 25 (55.8–201.1) | 32.9 | 30.6 | 25% | −15.2 | 0.03 | 0.17 | 69% 96% |
| MDRD 2$^f$          | 94.5 ± 25.0 (50.3–184.9) | 28.7 | 23.5 | 22% | −18.3 | 0.02 | 0.15 | 65% 95% |
| Jelliffe 1$^e$      | 92.3 ± 22.1 (55.3–176.0) | 26.0 | 26.5 | 23% | −13.2 | 0.05 | 0.23 | 67% 96% |
| Jelliffe 2          | 96.8 ± 22.2 (59.7–188.0) | 26.0 | 19.0 | 17% | −16.0 | 0.01 | 0.25 | 67% 95% |
| Mawer               | 141.6 ± 66.0 (40.9–396.7) | 49.6 | 35.4 | 30% | 28.7 | 0.07 | 0.26 | 50% 72% |
| Bjornsson           | 122.5 ± 37.3 (58.2–290.5) | 29.5 | 24.8 | 23% | 9.7 | 0.08 | 0.28 | 66% 86% |
| Gates               | 94.4 ± 24.9 (50.3–207.3) | 28.8 | 25.4 | 25% | −18.4 | 0.02 | 0.14 | 66% 95% |
| Corcoran-Salazar    | 114.5 ± 29.1 (56.7–199.4) | 26.1 | 20.9 | 19% | 1.7 | 0.04 | 0.21 | 76% 92% |

$^a$ CrCl, creatinine clearance; CG, Cockcroft-Gault; MDRD, Modification of Renal Disease.
$^b$ Mean calculated CrCl or GFR expressed as mean ± SD (range).
$^c$ Bias is the mean prediction error:
$$
\text{ME} = \frac{1}{N} \sum_{i=1}^{N} (\text{pe}_i - \text{true value})/N
$$

$^d$ $R^2$ statistic was derived by simple linear regression and reflects the predictive ability of the model.

$^e$ Adjusted for BSA of 1.73 m$^2$.
$^f$ Expressed as ml/min per 1.73 m$^2$. 

Copyright © American Society of Nephrology. Unauthorized reproduction of this article is prohibited.
especially critical when estimating GFR in subjects with normal or mildly impaired renal function because small changes in SCr result in large changes in calculated CrCl and GFR. For example, a 40-yr-old black woman who weighs 70 kg and has a SCr of 1.0 mg/dl in steady state has a calculated CrCl of 82.6 ml/min and a calculated GFR of 78.3 ml/min per 1.73 m²; an increase of SCr by 0.1 mg/dl to 1.1 mg/dl results in a calculated CrCl of 75.1 ml/min and a calculated GFR of 71.2 ml/min per 1.73 m². In contrast, the same patient with a SCr of 3.1 mg/dl in steady state would have an estimated CrCl of 27.5 ml/min and an estimated GFR of 26.1 ml/min per 1.73 m²; an increase in SCr to 3.2 mg/dl results in very small changes to 26.7 ml/min and 25.3 ml/min per 1.73 m² for CrCl and GFR, respectively. There is no standard calibration for SCr measurements for laboratories in the United States (22), and investigators have reported that SCr on the same stored serum samples were 0.23 mg/dl higher when assayed by the laboratory used by Third National Health and Nutrition Examination Survey study when compared with the values obtained by the laboratory used by the MDRD study, a magnitude of difference that was common across different laboratories (18).

Other potential sources of measurement error include (1) intraindividual variability in SCr; (2) intraindividual variability in other clinical or laboratory measurements; (3) intraindividual variability in GFR; and (4) intra-assay variability in GFR measurement. Despite the 125I-iothalamate protocol requirements for minimum urinary flow rates of 3 ml/min, this was not achieved in all nine subjects who had CV/H11350 15%. Intra-assay GFR measurement errors are likely minimized in those who underwent125 I-iothalamate testing with CV/H11349 10%; however, in those with high CV and those who underwent 99m Tc-DTPA studies without reported intra-assay CV, we acknowledge that the questionable quality of some of the GFR measurements may not make it a true and reliable “gold standard.”

<table>
<thead>
<tr>
<th></th>
<th>Mean CrCl/GFR(^a) (Range)</th>
<th>Mean Absolute Difference</th>
<th>Median Absolute Difference</th>
<th>Median % Absolute Difference</th>
<th>Bias(^b)</th>
<th>(R^2) (^c) (Precision)</th>
<th>Pearson Correlation</th>
<th>Accuracy % within 30%</th>
<th>Accuracy % within 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>(^{125})I-Iothalamate Measured</td>
<td>102.8 ± 15.7 (70–151.9)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>CG(^d)</td>
<td>108.3 ± 25.1 (57.8–169.4)</td>
<td>42.7</td>
<td>33.1</td>
<td>33%</td>
<td>25.6</td>
<td>0.04</td>
<td>0.13</td>
<td>45%</td>
<td>73%</td>
</tr>
<tr>
<td>CG-GFR(^d)</td>
<td>107.8 ± 42.1 (42.1–225.2)</td>
<td>33.8</td>
<td>28.6</td>
<td>25%</td>
<td>5.0</td>
<td>0.05</td>
<td>0.13</td>
<td>56%</td>
<td>78%</td>
</tr>
<tr>
<td>MDRD 1(^e)</td>
<td>102.3 ± 25.9 (55.8–201.2)</td>
<td>33.8</td>
<td>28.7</td>
<td>25%</td>
<td>-0.5</td>
<td>0.15</td>
<td>0.27</td>
<td>78%</td>
<td>96%</td>
</tr>
<tr>
<td>MDRD 2(^e)</td>
<td>99.4 ± 25.5 (50.3–184.9)</td>
<td>20.5</td>
<td>18.5</td>
<td>17%</td>
<td>-3.3</td>
<td>0.14</td>
<td>0.22</td>
<td>78%</td>
<td>95%</td>
</tr>
<tr>
<td>(^{99m})Tc-DTPA Measured</td>
<td>125.0 ± 20.3 (77–169)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>CG(^d)</td>
<td>106.5 ± 33.0 (63.8–259.2)</td>
<td>31.6</td>
<td>22.0</td>
<td>19%</td>
<td>6.0</td>
<td>0.18</td>
<td>0.41</td>
<td>73%</td>
<td>87%</td>
</tr>
<tr>
<td>CG-GFR(^d)</td>
<td>110.1 ± 40.0 (48.0–242.3)</td>
<td>31.7</td>
<td>31.1</td>
<td>25%</td>
<td>-15.0</td>
<td>0.21</td>
<td>0.42</td>
<td>62%</td>
<td>89%</td>
</tr>
<tr>
<td>MDRD 1(^e)</td>
<td>91.9 ± 22.8 (60.1–172.2)</td>
<td>31.8</td>
<td>31.1</td>
<td>25%</td>
<td>-33.2</td>
<td>0.40</td>
<td>0.45</td>
<td>58%</td>
<td>96%</td>
</tr>
<tr>
<td>MDRD 2(^e)</td>
<td>88.5 ± 23.2 (52.9–167.2)</td>
<td>38.8</td>
<td>38.2</td>
<td>32%</td>
<td>-36.5</td>
<td>0.38</td>
<td>0.43</td>
<td>49%</td>
<td>96%</td>
</tr>
</tbody>
</table>

\(^a\) Mean calculated CrCl or GFR expressed as mean ± SD (range).

\(^b\) Bias is the mean prediction error:

\[
\text{ME} = \sum \frac{p_e_i - t_i}{N}
\]

where \(p_e_i = \) predicted value - true value and \(N = \) sample size.

\(^c\) \(R^2\) statistic was derived by simple linear regression and reflects the predictive ability of the model.

\(^d\) Adjusted for BSA of 1.73 m².

\(^e\) Expressed as mL/min per 1.73 m².

Survey study when compared with the values obtained by the laboratory used by the MDRD study, a magnitude of difference that was common across different laboratories (18).
Table 4. Comparison of CG and MDRD equations in previously published and current studies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>1628</td>
<td>1703</td>
<td>22</td>
<td>109</td>
<td>92</td>
<td>26</td>
<td>100</td>
</tr>
<tr>
<td>Subjects</td>
<td>Patients with CKD in MDRD study</td>
<td>Blacks with CKD from hypertension</td>
<td>Potential renal donors with CrCl &lt;30 ml/min by two 24-h urines</td>
<td>CKD cohort</td>
<td>46 healthy adults/46 type 1 diabetics without nephropathy</td>
<td>46 healthy adults/46 type 1 diabetics without nephropathy</td>
<td>Potential kidney donors</td>
</tr>
<tr>
<td>Mean age (yr)</td>
<td>51</td>
<td>54</td>
<td>40</td>
<td>43</td>
<td>28/27</td>
<td>58</td>
<td>41</td>
</tr>
<tr>
<td>% male</td>
<td>60%</td>
<td>69%</td>
<td>27%</td>
<td>77%</td>
<td>46%</td>
<td>23%</td>
<td>42%</td>
</tr>
<tr>
<td>% black</td>
<td>12%</td>
<td>100%</td>
<td>9%</td>
<td>0%</td>
<td>0%</td>
<td>4%</td>
<td>10%</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>79.6 ± 16.8</td>
<td>90.2</td>
<td>71.4 ± 16.5</td>
<td>76</td>
<td>69.9/71.5</td>
<td>NR</td>
<td>78.1 ± 15.7</td>
</tr>
<tr>
<td>(range 66–85)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(range 44.1–113)</td>
</tr>
<tr>
<td>BSA (m²)</td>
<td>1.91 ± 0.23</td>
<td>2.02</td>
<td>1.80 ± 0.21</td>
<td>NR</td>
<td>1.85/1.87</td>
<td>NR</td>
<td>1.9 ± 0.2</td>
</tr>
<tr>
<td>Mean SCr (mg/dl)</td>
<td>2.3</td>
<td>1.85</td>
<td>1.1</td>
<td>1.2</td>
<td>0.90.8</td>
<td>NR</td>
<td>0.9</td>
</tr>
<tr>
<td>Measured GFR (ml/min per 1.73 m²)</td>
<td>48.6</td>
<td>56.9</td>
<td>99.3</td>
<td>100</td>
<td>107/122</td>
<td>NR</td>
<td>112</td>
</tr>
<tr>
<td>(range 18–205)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(range 70–169)</td>
</tr>
<tr>
<td>Type of GFR study</td>
<td>125I-Iothalamate</td>
<td>125I-Iothalamate</td>
<td>99mTc-DTPA Iohexol</td>
<td>Insulin</td>
<td>51Cr-EDTA</td>
<td>55 125I-Iothalamate/45 99mTc-DTPA</td>
<td></td>
</tr>
<tr>
<td>R or R² for MDRD 1</td>
<td>R² = 0.90 (log transformed)</td>
<td>R = 0.90 (log transformed)</td>
<td>R² = 0.005 (log transformed)</td>
<td>R² = 0.31 (log transformed)</td>
<td>NR</td>
<td>NR</td>
<td>R² = 0.15/0.40</td>
</tr>
<tr>
<td>R or R² for CG-GFR</td>
<td>R² = 0.80 (log transformed)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>R² = 0.71 (log transformed)</td>
<td>R² = 0.05/0.21</td>
</tr>
<tr>
<td>R or R² for CG</td>
<td>NR</td>
<td>R = 0.85 (log transformed)</td>
<td>R² = 0.14 (log transformed)</td>
<td>R² = 0.21 (log transformed)</td>
<td>NR</td>
<td>NR</td>
<td>R² = 0.04/0.18</td>
</tr>
<tr>
<td>Absolute difference for MDRD 1 (ml/min per 1.73 m²)</td>
<td>NR</td>
<td>6.27 (median)</td>
<td>NR</td>
<td>NR</td>
<td>9.0/11.8 (median)</td>
<td>NR</td>
<td>33.8/31.8 (median)</td>
</tr>
<tr>
<td>Absolute difference for CG (ml/min per 1.73 m²)</td>
<td>NR</td>
<td>8.34 (median)</td>
<td>NR</td>
<td>NR</td>
<td>10.7/18.8 (median)</td>
<td>NR</td>
<td>33.8/31.7 (median)</td>
</tr>
</tbody>
</table>

a BSA, body surface area; SCr, serum creatinine; NR, not reported.
the intercept terms are different for both MDRD 1 and MDRD 2, suggesting that, as expected, our study population varies considerably from the original cohort in which these equations were derived. Furthermore, the regression coefficients for all laboratory assays (Scr, BUN, and alb) are also considerably different, likely reflecting lack of calibration in a central laboratory and measurement error, whereas the regression coefficients for age and gender are similar. The coefficient for black race is in the opposite direction of those in the MDRD equations but are based on only 10 blacks. Using the MDRD predictors of GFR resulted in models that explain only 8 to 10% of the variance in our data set of normal subjects, which is not surprising in light of the relatively small sample size of 100 subjects and the various sources of measurement error as previously discussed, especially those involving Scr calibration.

Our results are consistent with other published investigations in the literature (Table 4). A study by Bostom et al. (6) of 109 patients with known kidney disease and Scr \( \leq 1.5 \text{ mg/dl} \) reported an \( R^2 \) of 0.31 for MDRD 1, 0.29 for MDRD 2, and 0.17 for CG (Table 4). The Bostom study differed from this current study on several points: it included patients with renal disease, there was a narrower range of weights among the subjects, the majority of subjects had some degree of proteinuria, and io-hexol was used to determine GFR (Table 4). Despite the “normal serum creatinine levels” (range, 1.0 to 1.3 mg/dl), measured GFR by io-hexol was as low as 18 ml/min per 1.73 m\(^2\) (25th to 75th percentile range was 88 to 138 ml/min per 1.73 m\(^2\)). In addition, 25% had io-hexol GFR <80 ml/min per 1.73 m\(^2\), and 59% had glomerular disease, whereas our cohort were healthy, normal adults.

Another recent study of 46 healthy adults and 46 individuals who had type 1 diabetes without evidence of nephropathy and underwent insulin clearance studies found median absolute differences of 10.7 versus 9.0, respectively, when comparing MDRD 1 and CG equations in healthy subjects; the CG equation was corrected to reflect GFR by either a multiplication factor of 0.9 and the formula \( Y = -0.004 + 1.54 \) (Table 4) (7). No analyses of bias, precision, or accuracy were presented (7). On the basis of the higher median absolute differences in the MDRD 1 equation compared with the CG equation for all subjects, the authors concluded that the MDRD 1 equation underestimated GFR, especially in women with diabetes, and that the MDRD 1 equation was less “accurate” than the CG equation. Whereas we also found that the MDRD 1 equation consistently underestimated GFR measurements (most notably in those who underwent \(^{99m}\)Tc-DTPA), we observed that the MDRD equations were less biased, more precise, and more accurate than the CG equation.

Several limitations of this study should be noted. First, very few black individuals were included, so the findings may not be generalizable to this group. Second, these analyses were performed on a relatively small sample size with wide variability in clinical and laboratory parameters. Our study participants, however, are likely representative of the general population with normal or mildly decreased kidney function. Third, Scr measurements were not performed at the same time in the same laboratory and were not calibrated with a standardized measurement, although the vast majority (82%) of those who underwent \(^{125}\)I-iothalamate studies had Scr measured in the same BWH laboratory with a low reported CV. Despite these limitations, however, this is the first study to evaluate rigorously these GFR formulas in healthy individuals without chronic kidney disease. Moreover, the lack of calibration of Scr across different laboratories continues to be a reality that physicians face daily in clinical practice.

In addition to the inaccuracies of the formulas themselves, this study underscores how multiple sources of measurement error (including intra-assay Scr variability, intra-individual Scr variability, lack of calibration of Scr assays across different laboratories, intra-assay GFR variability, intra-individual GFR variability, and measurement error of other variables in the prediction equations) can affect the precision and accuracy of renal clearance prediction equations. In the future, researchers and clinicians would greatly benefit from a study of a large and racially diverse cohort of people with normal or mildly impaired kidney function that would allow investigators to derive an improved prediction equation for estimating GFR.

Acknowledgment

This work was supported in part by the DCI Paul Teschan Research Fund.

References


Access to UpToDate on-line is available for additional clinical information at http://www.jasn.org/
Correction


A reader has correctly pointed out that we made an error in standardizing the Cockcroft-Gault (CG) and Jelliffe-1 equations for a body surface area (BSA) of 1.73 m². We multiplied these equations by BSA/1.73 m² when we should have multiplied by 1.73 m²/BSA. We submit the highlighted corrected data in Table 1 below alongside the previously reported values for comparison.

All other previously reported data, including those for the MDRD equations, are not affected by this error. We must therefore revise some of our original conclusions. The CG equations also appear to underestimate measured GFR in adults without kidney disease, and the accuracy of the CG equations are comparable to those of the MDRD equations. These data do still show, however, that estimation equations are less than optimal for evaluating GFR in healthy adults with serum creatinine in the normal range.

### Table 1. Mean calculated CrCl or GFR, mean absolute difference, bias, precision, and accuracy of GFR prediction equations compared to measured GFR

<table>
<thead>
<tr>
<th></th>
<th>Mean CrCl/GFR (range)</th>
<th>Mean Absolute Difference</th>
<th>Median Absolute Difference</th>
<th>Median % Absolute Difference</th>
<th>Bias</th>
<th>R² (Precision)</th>
<th>Pearson Correlation</th>
<th>Accuracy % within:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50%</td>
</tr>
<tr>
<td>CG</td>
<td>129.6 ± 48.8 (50.1–288.5)</td>
<td>37.9</td>
<td>26.7</td>
<td>23%</td>
<td>16.8</td>
<td>0.06</td>
<td>0.24</td>
<td>58%</td>
</tr>
<tr>
<td>CORRECTED CG</td>
<td>107.5 ± 28.9 (57.8–259.2)</td>
<td>24.5</td>
<td>21.5</td>
<td>19%</td>
<td>−5.3</td>
<td>0.09</td>
<td>0.30</td>
<td>78%</td>
</tr>
<tr>
<td>CG-FGR</td>
<td>108.9 ± 41 (42.1–242.3)</td>
<td>32.9</td>
<td>30.5</td>
<td>25%</td>
<td>−4.0</td>
<td>0.06</td>
<td>0.25</td>
<td>59%</td>
</tr>
<tr>
<td>CORRECTED CG-FGR</td>
<td>96.3 ± 24.2 (48.6–217.7)</td>
<td>28.5</td>
<td>27.5</td>
<td>24%</td>
<td>−22.5</td>
<td>0.09</td>
<td>0.30</td>
<td>61%</td>
</tr>
<tr>
<td>MDRD 1</td>
<td>97.6 ± 25 (55.8–201.1)</td>
<td>32.9</td>
<td>30.6</td>
<td>25%</td>
<td>−15.2</td>
<td>0.03</td>
<td>0.17</td>
<td>69%</td>
</tr>
<tr>
<td>MDRD 2</td>
<td>94.5 ± 25 (50.3–184.9)</td>
<td>28.7</td>
<td>23.5</td>
<td>22%</td>
<td>−18.3</td>
<td>0.02</td>
<td>0.15</td>
<td>65%</td>
</tr>
<tr>
<td>Jelliffe 1</td>
<td>92.3 ± 22.1 (55.3–176.0)</td>
<td>26.0</td>
<td>26.5</td>
<td>23%</td>
<td>−13.2</td>
<td>0.05</td>
<td>0.23</td>
<td>67%</td>
</tr>
<tr>
<td>CORRECTED JELLIF-E-1</td>
<td>86.5 ± 25.1 (80.3–166.8)</td>
<td>26.0</td>
<td>26.9</td>
<td>26%</td>
<td>−26.3</td>
<td>0.03</td>
<td>0.16</td>
<td>61%</td>
</tr>
</tbody>
</table>

Estimation equations:

1. Cockcroft-Gault (CG): CrCl = [(140 – age) × weight (kg)/SCr × 72] × 0.85 (if female)
2. Cockcroft-Gault FGR (CG-FGR) estimate: GFR = 0.84 × CrCl by Equation (1)
3. MDRD 1: GFR = 170 × (SCr)^-0.990 × (age)^-0.170 × (0.762 if female) × (1.18 if black) × (BUN)^-0.170 × (Alb)^0.218
4. MDRD 2: GFR = 186 × (SCr)^-1.154 × (age)^-0.203 × (0.742 if female) × (1.212 if black)
5. Jelliffe 1: (1.73 m²/BSA): (98 – [0.8 × (age – 20)])/SCr × [0.90 if female])

CrCl, creatinine clearance; BSA, body surface area; SCr, serum creatinine; MDRD, Modification of Diet in Renal Disease; BUN, blood urea nitrogen; Alb, albumin.

### References


These findings are supported by other recent publications. Rule *et al.* analyzed 274 potential kidney donors (serum creatinine, 0.7 to 1.6 mg/dl) who underwent iothalamate clearance studies and found a bias of −29 ml/min per 1.73 m² for the MDRD2 equation and a bias of −14 ml/min per 1.73 m² for the CG equation standardized to BSA. Correlation coefficients were 0.26 and 0.30 for MDRD2 and CG, respectively (1). Poggio *et al.* analyzed 457 potential kidney donors with iothalamate clearance and reported a median difference of −9.0 ml/min per 1.73 m² and a median absolute difference of 16 ml/min per 1.73 m² for GFR estimates by the MDRD2 (2). These two studies concluded, as did ours, that the MDRD2 equation significantly underestimates measured GFR in adults without kidney disease, and that a formula developed in a cohort with moderate to severe chronic kidney disease understandably does not perform well in adults without chronic kidney disease. Because estimation equations are essential in large cohort studies and are important in clinical settings such as kidney donation, we advocate further study and the development of more precise, more accurate, and less biased prediction equations for adults with serum creatinine in the normal range.

Julie Lin, MD, MPH; Eric L. Knight, MD, MPH; Ajay Singh, MD