New Equations to Estimate GFR in Children with CKD

George J. Schwartz,* Alvaro Muñoz,† Michael F. Schneider,† Robert H. Mak,‡ Frederick Kaskel,§ Bradley A. Warady,† and Susan L. Furth†¶

*Department of Pediatrics, University of Rochester School of Medicine, Rochester, and §Department of Pediatrics, Albert Einstein College of Medicine, Bronx, New York; †Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health and ¶Department of Pediatrics, Johns Hopkins School of Medicine, Baltimore, Maryland; ‡Department of Pediatrics, Children’s Hospital of San Diego, San Diego, California; †Department of Pediatrics, Children’s Mercy Hospital, Kansas City, Missouri

ABSTRACT
The Schwartz formula was devised in the mid-1970s to estimate GFR in children. Recent data suggest that this formula currently overestimates GFR as measured by plasma disappearance of iohexol, likely a result of a change in methods used to measure creatinine. Here, we developed equations to estimate GFR using data from the baseline visits of 349 children (aged 1 to 16 yr) in the Chronic Kidney Disease in Children (CKiD) cohort. Median iohexol-GFR (iGFR) was 41.3 ml/min per 1.73 m² (interquartile range 32.0 to 51.7), and median serum creatinine was 1.3 mg/dl. We performed linear regression analyses assessing precision, goodness of fit, and accuracy to develop improvements in the GFR estimating formula, which was based on height, serum creatinine, cystatin C, blood urea nitrogen, and gender. The best equation was

\[
GFR = 39.1 \times \frac{\text{height (m)}}{\text{Scr (mg/dl)}}^{0.516} \times \left[\frac{1.8}{\text{cystatin C (mg/L)}}\right]^{0.294} \times \left[\frac{30}{\text{BUN (mg/dl)}}\right]^{0.169}[1.099]^{\text{male}} \times \frac{\text{height (m)}}{1.4}^{0.188}.
\]

This formula yielded 87.7% of estimated GFR within 30% of the iGFR, and 45.6% within 10%. In a test set of 168 CKiD patients at 1 yr of follow-up, this formula compared favorably with previously published estimating equations for children. Furthermore, with height measured in cm, a bedside calculation of 0.413*(height/serum creatinine), provides a good approximation to the estimated GFR formula. Additional studies of children with higher GFR are needed to validate these formulas for use in screening all children for CKD.


GFR is the most useful indicator of kidney function and kidney disease progression; however, determination of true GFR is time-consuming, costly, and difficult to perform for regular clinical use in children. Thus, there is considerable interest in developing formulas to estimate GFR using endogenous surrogate markers such as creatinine1,2 or the low molecular weight protein cystatin C.3–5 The Schwartz formula, devised for children in the mid-1970s,1,6 estimates GFR from an equation that uses serum creatinine (Scr), height, and an empirical constant.1,2,6 GFR as estimated by the Schwartz formula has been used as one of the enrollment criteria for the Chronic Kidney Disease in Children (CKiD) cohort.

Received March 12, 2008. Accepted August 26, 2008.

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

Correspondence: Dr. George J. Schwartz, Professor of Pediatrics and of Medicine, Chief, Pediatric Nephrology, Box 777, University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, NY 14642. Phone: 585-275-9784; Fax: 585-756-8054; E-mail: george_schwartz@urmc.rochester.edu

Copyright © 2009 by the American Society of Nephrology
Markers of kidney function

- **CKD cause (%)**
  - Non-GN: 73
  - GN: 20
  - IV: 14
  - other: 7
- **Tanner stage (%)**
  - I: 60
  - II: 11
  - III: 8
  - IV: 14
  - V: 7
- **Hispanic ethnicity (%)**: 15
- **Male gender (%)**: 61
- **Age (yr; median)**: 0.621 (0.7 to 1.4)
- **Height (m; median)**: 0.102 49.0 0.075
- **Weight (kg; median)**: 0.060 18.9 0.510
- **Serum creatinine (mg/dl; median)**: 0.042 38.7
- **Cystatin C (mg/L; median)**: 0.034 0.2
- **BUN (mg/dl; median)**: 0.034 39.0
- **Iohexol GFR (ml/min per 1.73 m²; median)**: 0.033 44.4

Markers of body size

- **Male gender (%)**: 61
- **Race (%)**
  - white: 69
  - black: 15
  - other: 8
  - >1 race: 8
- **Hispanic ethnicity (%)**: 15
- **Height percentile (median)**: 1.870 (1.73 to 2.03)
- **Height (m; median)**: 0.102 49.0 0.075
- **Weight percentile (median)**: 0.067 1.2 0.545
- **Weight (kg; median)**: 0.034 39.0
- **BMI (kg/m²; median)**: 0.060 18.9 0.510
- **BMI percentile (median)**: 0.033 44.4
- **BSA (m²; median)**: 0.033 44.4
- **BSA (m²; median)**: 0.060 18.9 0.510

**BSA-Unadjusted iGFR**

- **Regression Coefficient ± SE**: 0.034 ± 0.059
- **R² (%)**: 0.1

**BSA-Adjusted iGFR**

- **Regression Coefficient ± SE**: 0.035 ± 0.039
- **R² (%)**: 0.2

**Markers of body size**

- **Age (yr)**: 0.621 ± 0.042
- **Height (m)**: 1.870 ± 0.102
- **BSA (m²)**: 1.074 ± 0.050

**Markers of kidney function**

- **1/serum creatinine (mg/dl)**: 0.138 ± 0.066
- **1/cystatin C (mg/L)**: 0.745 ± 0.083
- **1/BUN (mg/dl)**: 0.539 ± 0.060

**Dependent Variable**

<table>
<thead>
<tr>
<th>Variables</th>
<th>BSA-Unadjusted iGFR</th>
<th>BSA-Adjusted iGFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>markers of body size</td>
<td>Regression Coefficient ± SE</td>
<td>R² (%)</td>
</tr>
<tr>
<td>marker of gender</td>
<td>0.034 ± 0.059</td>
<td>0.1</td>
</tr>
<tr>
<td>marker of age</td>
<td>0.621 ± 0.042</td>
<td>38.7</td>
</tr>
<tr>
<td>marker of weight</td>
<td>0.725 ± 0.034</td>
<td>57.4</td>
</tr>
<tr>
<td>marker of height</td>
<td>1.870 ± 0.102</td>
<td>49.0</td>
</tr>
<tr>
<td>marker of BSA</td>
<td>1.074 ± 0.050</td>
<td>56.9</td>
</tr>
</tbody>
</table>

**Markers of kidney function**

- **1/serum creatinine (mg/dl)**: 0.138 ± 0.066
- **1/cystatin C (mg/L)**: 0.745 ± 0.083
- **1/BUN (mg/dl)**: 0.539 ± 0.060

**Table 2.** Values from univariate linear regression models used to determine the predictability of BSA-unadjusted and BSA-adjusted iGFR for 349 children

**Table 1.** Baseline characteristics of 349 children of the CKID study

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male gender (%)</td>
<td>61</td>
</tr>
<tr>
<td>Race (%)</td>
<td>69 - white, 15 - black, 8 - other, 8 - &gt;1 race</td>
</tr>
<tr>
<td>Hispanic ethnicity (%)</td>
<td>15</td>
</tr>
<tr>
<td>Male gender (%)</td>
<td>61</td>
</tr>
<tr>
<td>Age (yr; median)</td>
<td>10.8 (7.7 to 14.3)</td>
</tr>
<tr>
<td>Height (m; median)</td>
<td>1.4 (1.2 to 1.6)</td>
</tr>
<tr>
<td>Height percentile (median)</td>
<td>22.8 (5.5 to 51.3)</td>
</tr>
<tr>
<td>Weight (kg; median)</td>
<td>35.0 (23.6 to 53.0)</td>
</tr>
<tr>
<td>Weight percentile (median)</td>
<td>45.3 (15.0 to 78.2)</td>
</tr>
<tr>
<td>BSA (m²; median)</td>
<td>1.2 (0.9 to 1.5)</td>
</tr>
<tr>
<td>BMI (kg/m²; median)</td>
<td>18.1 (16.2 to 21.6)</td>
</tr>
<tr>
<td>BMI percentile (median)</td>
<td>64.4 (35.3 to 88.4)</td>
</tr>
<tr>
<td>Serum creatinine (mg/dl; median)</td>
<td>1.3 (1.0 to 1.8)</td>
</tr>
<tr>
<td>Height (m)/serum creatinine (mg/dl; median)</td>
<td>1.0 (0.8 to 1.3)</td>
</tr>
<tr>
<td>Cystatin C (mg/L; median)</td>
<td>1.8 (1.5 to 2.3)</td>
</tr>
<tr>
<td>BUN (mg/dl; median)</td>
<td>27.0 (22.0 to 38.0)</td>
</tr>
<tr>
<td>Iohexol GFR (ml/min per 1.73 m²; median)</td>
<td>41.3 (32.0 to 51.7)</td>
</tr>
</tbody>
</table>

**BSA-Unadjusted iGFR**: from univariate linear regression models used to determine the predictability of BSA-unadjusted and BSA-adjusted iGFR for 349 children.

**Table 2.** Values from univariate linear regression models used to determine the predictability of BSA-unadjusted and BSA-adjusted iGFR for 349 children

- **Variables**
  - markers of body size: gender, age, height, BSA
  - markers of kidney function: serum creatinine, cystatin C, BUN
- **Dependent Variable**
  - BSA-Unadjusted iGFR: regression coefficient ± SE, R² (%)
  - BSA-Adjusted iGFR: regression coefficient ± SE, R² (%)

- **BSA-Unadjusted iGFR**
  - marker of gender: 0.034 ± 0.059, R² = 0.1
  - marker of age: 0.621 ± 0.042, R² = 38.7
  - marker of height: 1.870 ± 0.102, R² = 49.0
  - marker of BSA: 1.074 ± 0.050, R² = 56.9

- **BSA-Adjusted iGFR**
  - marker of gender: 0.035 ± 0.039, R² = 0.2
  - marker of age: -0.007 ± 0.035, R² < 0.1
  - marker of height: 0.075 ± 0.094, R² = 0.2
  - marker of BSA: 0.074 ± 0.050, R² = 0.6

- **Markers of body size**
  - gender: 0.034 ± 0.059, R² = 0.1
  - age: 0.621 ± 0.042, R² = 38.7
  - height: 1.870 ± 0.102, R² = 49.0
  - BSA: 1.074 ± 0.050, R² = 56.9

- **Markers of kidney function**
  - serum creatinine: 0.138 ± 0.066, R² = 1.2
  - cystatin C: 0.745 ± 0.083, R² = 18.9
  - BUN: 0.539 ± 0.060, R² = 18.9
RESULTS

Characteristics of Study Population

Of the 349 children studied, 61% were male, 69% were white, 15% were black, 79% were Tanner stages I through III, and only 20% had a form of glomerulonephritis (Table 1). The median age was 10.8 yr. Body habitus showed notable growth retardation, in that the median age- and gender-specific height percentile was 22.8 compared with the median age- and gender-specific weight percentile of 45.3. The median values of biochemical predictors of kidney function were 1.3 mg/dl, 1.8 mg/L, and 27 mg/dl for Scr, cystatin C, and BUN, respectively. The median iGFR was 41.3 ml/min per 1.73 m² with an interquartile range from 32.0 to 51.7. Ninety-five percent of the iGFR values were between 21.1 and 75.9 ml/min per 1.73 m², indicating that the distribution was positively skewed. Fewer than 10% had nephrotic-range proteinuria (urine protein-to-creatinine ratio >2.0), and the median serum albumin was 3.7: none had nephrotic syndrome.

Univariate Linear Regression Analyses

We performed univariate analyses of body surface area (BSA)-unadjusted iGFR on markers of body size and biochemical markers of kidney function (Table 2). BSA and weight had the highest $R^2$ values (57.4 and 56.9%, respectively). Furthermore, the regression coefficient for log(BSA) was 1.074 (not statistically different from 1, $P = 0.140$); therefore, the classical calibration to a BSA of 1.73 m² corresponds to the residuals of the regression, and adjusting iGFR for BSA essentially removed the variability in GFR that was attributable to the high variation in body size in our pediatric population. After adjustment for BSA, none of the body size variables had any additional predictive power, and the ability of the endogenous kidney markers to explain the variability of individuals of similar body sizes substantially increased (Table 2). Specifically, height/Scr, as previously shown by Schwartz et al., explained the greatest proportion of the variability ($R^2 = 65.0\%$) in iGFR when compared with the reciprocals of Scr ($R^2 = 44.4\%$), cystatin C ($R^2 = 47.3\%$), and BUN ($R^2 = 39.0\%$).

Figures 1 through 3 show the scatter plots (in logarithmic scale) of iGFR versus height/Scr, 1.8/cystatin C, and 30/BUN, respectively. The relationships between iGFR and each of these biochemical markers were appropriately described by regression lines, because there was good agreement with nonparametric splines depicted by dashed curves in the figures.

Model-Based eGFR Formulas

Table 3 shows the regression analyses using the overall mean as the estimate for all individuals (i.e., no model) to the model using height/Scr, cystatin C, BUN, gender, and height (model III). When no model was used, the square root of the mean square error (0.351) was simply the SD of the 349 eGFR values in the logarithmic scale. The updated Schwartz formula corresponds to the particular case of imposing the exponent of height/Scr to be 1, resulting in the equation

$$eGFR = 41.3[\text{height}/\text{Scr}],$$

which yielded 79.4% of estimated GFR values within 30% of the measured iGFR. If height were reported in cm instead of meters, then this updated Schwartz formula would be

$$eGFR = 0.413[\text{height}/\text{Scr}],$$

showing a 25% reduction from the previous 0.55 generated by the Jaffe-based Scr measurements, in keeping with the approximate reduction in apparent concentration by isotope dilution mass spectroscopy—referred enzymatic creatinine determinations.1,6 Figure 1 indicates that the exponent of 1 of height/SCR

*Figure 1. Analysis of log-transformed height/Scr and iGFR, showing that 65.0% of the variation in log(iGFR) can be explained by log(height/Scr). Regression line and nonparametric spline depicted by dashed curve are superimposed.*

*Figure 2. Analysis of log-transformed 1.8/cystatin C and iGFR, showing that 47.3% of the variation in log(iGFR) can be explained by the reciprocal of cystatin C. Regression line and nonparametric spline depicted by dashed curve are superimposed.*
Spondes to the case of the shape parameter being equal to 0.19 ing appropriate to describe the distribution of iGFR as it corre-

interval

found that the shape parameter was

log(iGFR) as a Gaussian variate], we allowed model III to have

assess the goodness of fit of the lognormal model [within 30 and 10%, respectively, of the measured iGFR. To

0.176, resulting in 87.7 and 45.6% of the eGFR values falling

R

metric spline depicted by dashed curve are superimposed.

reciprocal of BUN concentration. Regression line and nonpara-

that 39.0% of the variation in log(iGFR) can be explained by the

Scrn in the updated Schwartz formula is not correct, because the

estimate of the exponent was 0.775, significantly lower than 1. Models IA, IB, and IC in Table 3 show the three bivariate mod-
els with each of the three pairs of biochemical markers. Adding

cystatin C or BUN to a model with height/Scr improved the $R^2$
to approximately 69%. Model IC, which included only cystatin
C and BUN, did not perform as well. When all three variables

were incorporated into model II, root mean square error de-
greater to 0.185. We then tested whether there was any addi-
tional predictive power of gender, age, weight, height, BSA, Tanner stage, race, and body mass index. We found that the

addition of gender and height (alone) significantly improved the eGFR. This model III,

eGFR = 39.1[height/Scr]$^{0.516}$[1.8/cystatin C]$^{0.294}$

$\times [30/BUN]^{0.169}[1.099]^{male}[height/1.4]^{0.188}$,

showed $R^2$ of 75.2% with root mean square error down to

0.176, resulting in 87.7 and 45.6% of the eGFR values falling

within 30 and 10%, respectively, of the measured iGFR. To

assess the goodness of fit of the lognormal model [i.e., log(iGFR) as a Gaussian variate], we allowed model III to have

residual error distributed as a generalized gamma variate and

found that the shape parameter was $-0.0098$ (95% confidence interval $-0.23$ to 0.21), consistent with the lognormal model be-
ing appropriate to describe the distribution of iGFR as it corre-
sponds to the case of the shape parameter being equal to 0.19

Comparison with Other GFR Prediction Equations

We examined two creatinine-based, two cystatin C–based, and
two creatinine- and cystatin C–based prediction equations us-
ing published coefficients as well as coefficients derived from the

CKiD training data set of 349 children (Table 4). In general,

the model coefficients of the previously published formulas

were different from those obtained using the CKiD data. This

could be due to differing methods of measuring GFR, creati-
nine, or cystatin C or to CKiD’s focus on children with lower

levels of GFR.

Application to the Visit 2 Testing Data Set

All formulas shown in Table 4 and model III from Table 3 were

used to obtain eGFR for the 168 participants whose iGFR was

again measured at visit 2, which was scheduled to occur 1 yr

after the baseline visit. Table 5 shows the mean and SD of eGFR

as well as the bias, 95% limits of agreement, correlation, and

the percentage of eGFR values within 30 and 10% of the mea-

sured iGFR values.

The Counahan$^2$ and updated Schwartz formulas performed

comparably and adequately, with absolute bias $<2$ ml/min per

1.73 m$^2$ and correlation of 0.84. The creatinine- and cystatin
C–based formulas outperformed the single endogenous

marker formulas, especially when using the coefficients based

on the CKiD data. The Bouvet$^{20}$ and Zapitelli$^{21}$ formulas based

on the CKiD data had an absolute bias of approximately 2

ml/min per 1.73 m$^2$ and correlation of 0.86, and approximately

81 and 38% of eGFR values were within 30 and 10%, respect-
ively, of iGFR values; however, using the published coeffi-
cients,$^{20,21}$ there was more imprecision and less accuracy.

When model III was applied, the precision was better (limits of

agreement range under 30), and 83 and 41% of eGFR values

were within 30 and 10%, respectively, of measured iGFR val-

ues. Figure 4 depicts the Bland-Altman plot of eGFR values

using model III and iGFR values showing a strong correlation

($r = 0.88$) with a small bias of $-2.23$ ml/min per 1.73 m$^2$ and a

significantly lower SD of the eGFR values, as they correspond to

the mean values of the GFR for a given constellation of the

predictors (i.e., the eGFR values do not incorporate the error of

the regression coefficients or the residual error of the regres-
sion model).

DISCUSSION

Generation of formulas using endogenous serum substrates to

estimate GFR is useful in clinical medicine in view of the need
to adjust the dosage of nephrotoxic medication in the setting of

CKD to prevent further kidney damage. Furthermore, in the

CKiD study, GFR is measured directly using disappearance

curves of iohexol at the first visit and at all even visits (2, 4, 6,

etc.), but endogenous biochemical markers are measured at all

visits. The use of the estimating equations at the odd visits will

provide a means to have comparable data on GFR at all visits

and thus increase the power of the study to describe the trajec-
tories of the GFR decline. Indeed, well-established methods of mul-
tiple imputation should be implemented to account appropri-
ately for the imprecision of the estimation at the odd visits.

Endogenous creatinine clearance has been widely used to

measure GFR, but this measurement is affected by inaccuracies

in quantitatively collecting urine and by the renal tubular se-

cretion of creatinine, which would falsely elevate the apparent
Moreover, there are methodologic interferences in the measurement of the true concentration of creatinine, and there is a lack of availability of pediatric creatinine serum standards referenced to an isotope dilution mass spectrometry method. In the CKID study, the largest prospective cohort study of CKD in children in North America, we have generated a new eGFR formula, based on an enzymatic creatinine method. The most widely used estimate of GFR is the original Schwartz formula, which was generated from a highly significant correlation between GFR and $k \cdot \text{height/Scr}$. The current analysis of all of the variables in the CKID population shows that height/Scr still provides the best correlation with iGFR ($R^2 = 65.0\%$; see Figure 1), indicating that a parameter of body habitus along with Scr provides a useful measure of kidney function. Whereas the reciprocal of cystatin C showed a somewhat weaker correlation with iGFR ($R^2 = 47.3\%$; see Figure 2), the addition of both cystatin C and BUN to the height/Scr equation (model II) substantially improved the eGFR (Table 3). When all three variables were incorporated into a gender-based equation with an added coefficient for height alone (model III), there was further improvement in root mean square error (0.176) and $R^2 (75.2\%)$, and 88 and 46% of the estimates fell within 30 and 10%, respectively, of iGFR values, which is quite comparable to the best equations developed for adults; however, previous studies showed that the adult GFR estimating formulas derived from Cockcroft-Gault and the Modification of Diet in Renal Disease (MDRD) are not useful for children.

Two recent studies estimated GFR in children with higher GFR using equations including both Scr and cystatin C. We subjected the CKID data to the models reported by Bouvet et al. and Zappitelli et al. Whereas these equations with their published coefficients did not perform optimally in our test data set, after correction of their coefficients to the CKID data, their precision and accuracy were only slightly lower than our model III formula (Table 5). We cannot explain why their original equations did not perform well in our test data set, but differences in GFR level and methods of measuring GFR, cystatin C, and creatinine probably necessitated correcting their coefficients to the CKID data to optimize performance.

The formulas proposed here to estimate GFR using easily obtained biochemical markers have been developed in a group of children with mild to moderate CKD. A number of limitations preclude rapid generalization of this formula to the general pediatric population for estimation of GFR. Our population with moderate CKD has a median height percentile of 22.8%. Although we collected no direct measures of muscle mass in the CKID study, we have evidence of delayed puberty compared with normal children. The relationship between eGFR and the biochemical markers may be different in this population than in a population with more normal kidney function and without poor skeletal growth.

Similarly, although we did not observe in our population a change in formula with puberty, other populations with more normal body habitus should be examined to evaluate eGFR coef-
Table 4. Comparison of GFR prediction equations using published coefficients and coefficients derived from training data set of 349 children of the CKiD studya

<table>
<thead>
<tr>
<th>Model</th>
<th>Data Source</th>
<th>N</th>
<th>Ageb</th>
<th>Agent</th>
<th>GFRb</th>
<th>Equation for eGFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scr-based formulas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Counahan et al.</td>
<td>Original</td>
<td>108</td>
<td>0.2</td>
<td>51Cr-EDTA</td>
<td>4 to 200</td>
<td>43.00[height/Scr]</td>
</tr>
<tr>
<td>(n = 103)</td>
<td>CKID</td>
<td>349</td>
<td>1.2</td>
<td>lohexol</td>
<td>16 to 93</td>
<td>41.30[height/Scr]</td>
</tr>
<tr>
<td>Leger et al.35</td>
<td>Original</td>
<td>97</td>
<td>0.8</td>
<td>51Cr-EDTA</td>
<td>97 (31 to 200)</td>
<td>0.641[weight/Scr] + 16.063[height²/Scr]</td>
</tr>
<tr>
<td>CKID</td>
<td>349</td>
<td>1.2</td>
<td>lohexol</td>
<td>44 (16 to 93)</td>
<td>0.542[weight/Scr] + 9.948[height²/Scr]</td>
<td></td>
</tr>
<tr>
<td>Cystatin C–based formulas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Filler et al.36</td>
<td>Original</td>
<td>536</td>
<td>11.2</td>
<td>99mTc-DTPA</td>
<td>103 ± 41</td>
<td>91.62[1/CysC]² ¯1.23</td>
</tr>
<tr>
<td>CKID</td>
<td>349</td>
<td>10.6</td>
<td>lohexol</td>
<td>44 ± 15</td>
<td>66.22[1/CysC]³ ¯0.77</td>
<td></td>
</tr>
<tr>
<td>Grubb et al.27</td>
<td>Original</td>
<td>536</td>
<td>&lt;18</td>
<td>lohexol</td>
<td>113, 99, 63a</td>
<td>84.69[1/CysC]³ ¯0.60(1.384[age &lt; 14])</td>
</tr>
<tr>
<td>CKID</td>
<td>349</td>
<td>&lt;18</td>
<td></td>
<td>41, 42, NAa</td>
<td>68.06[1/CysC]³ ¯0.71(0.966[age &lt; 14])</td>
<td></td>
</tr>
<tr>
<td>Scr- and cystatin C–based formulas</td>
<td>Original</td>
<td>100</td>
<td>1.4</td>
<td>51Cr-EDTA</td>
<td>95 (18 to 200)</td>
<td>63.2[1.2/CysC]³ ¯0.56[96/88.84]/(Scr)² ¯0.35[weight/45]³ ¯0.35[age/14]³ ¯0.40</td>
</tr>
<tr>
<td>CKID</td>
<td>349</td>
<td>1.2</td>
<td>lohexol</td>
<td>44 (16 to 93)</td>
<td>46.1[1.2/CysC]³ ¯0.55[weight/45]³ ¯0.77[age/14]³ ¯0.17</td>
<td></td>
</tr>
<tr>
<td>Zappitelli et al.21</td>
<td>Original</td>
<td>103</td>
<td>12.7</td>
<td>lothalamate</td>
<td>74 ± 36</td>
<td>43.82[1/CysC]³ ¯0.63[1/Scr]³ ¯0.54[1.35[height]]</td>
</tr>
<tr>
<td>CKID</td>
<td>349</td>
<td>10.6</td>
<td>lohexol</td>
<td>44 ± 15</td>
<td>25.38[1/CysC]³ ¯0.33[1/Scr]³ ¯0.60[1.88[height]]</td>
<td></td>
</tr>
</tbody>
</table>

aGFR, ml/min per 1.73 m²; eGFR, ml/min per 1.73 m²; except Leger’s and Bouvet’s in ml/min; height, m; Scr, mg/dl; CysC, cystatin C, mg/L; weight, kg.
bMedians for age <14, 14 to <18, ≥18 yr.

cSources from the rest of the population (data not shown).

Table 5. Application of published and CKiD prediction equations to testing data set of 168 children of the CKiD study at first follow-up visit with iGFR 40.47 ± 16.10 ml/min per 1.73 m²

<table>
<thead>
<tr>
<th>Model</th>
<th>Data Source</th>
<th>eGFRa</th>
<th>Biasb</th>
<th>95% LOAc</th>
<th>Correlation</th>
<th>% of eGFR within 30% of iGFR</th>
<th>% of eGFR within 10% of iGFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scr-based formulas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Counahan et al.</td>
<td>Original</td>
<td>40.32</td>
<td>−0.15</td>
<td>−18.5</td>
<td>0.84</td>
<td>71</td>
<td>32</td>
</tr>
<tr>
<td>(n = 103)</td>
<td>CKID</td>
<td>38.73</td>
<td>−1.75</td>
<td>−19.7</td>
<td>0.84</td>
<td>73</td>
<td>26</td>
</tr>
<tr>
<td>Leger et al.35</td>
<td>Original</td>
<td>54.37</td>
<td>13.90</td>
<td>−17.7</td>
<td>0.79</td>
<td>45</td>
<td>13</td>
</tr>
<tr>
<td>CKID</td>
<td>38.95</td>
<td>−1.52</td>
<td>−23.3</td>
<td>0.79</td>
<td>71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cystatin C–based formulas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Filler et al.36</td>
<td>Original</td>
<td>42.46</td>
<td>1.99</td>
<td>−20.6</td>
<td>0.78</td>
<td>65</td>
<td>26</td>
</tr>
<tr>
<td>CKID</td>
<td>38.18</td>
<td>−2.30</td>
<td>−22.5</td>
<td>0.77</td>
<td>72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grubb et al.27</td>
<td>Original</td>
<td>35.45</td>
<td>−5.02</td>
<td>−33.0</td>
<td>0.78</td>
<td>43</td>
<td>14</td>
</tr>
<tr>
<td>CKID</td>
<td>38.28</td>
<td>−2.19</td>
<td>−22.6</td>
<td>0.76</td>
<td>72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scr and cystatin C–based formulas</td>
<td>Original</td>
<td>51.50</td>
<td>11.03</td>
<td>−11.1</td>
<td>0.80</td>
<td>48</td>
<td>15</td>
</tr>
<tr>
<td>CKID</td>
<td>38.31</td>
<td>−2.17</td>
<td>−18.1</td>
<td>0.86</td>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zappitelli et al.21</td>
<td>Original</td>
<td>34.86</td>
<td>−5.61</td>
<td>−22.3</td>
<td>0.86</td>
<td>65</td>
<td>32</td>
</tr>
<tr>
<td>CKID</td>
<td>38.27</td>
<td>−2.20</td>
<td>−17.8</td>
<td>0.87</td>
<td>82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CKiD equation IIIa</td>
<td>CKID</td>
<td>38.25</td>
<td>−2.23</td>
<td>−17.1</td>
<td>0.88</td>
<td>83</td>
<td>41</td>
</tr>
</tbody>
</table>

aGFR, ml/min per 1.73 m²; LOA, limits of agreement.
bBias = average of 168 (eGFR – iGFR) values, in ml/min per 1.73 m².
c95% LOA = bias ± 1.96 × SD of (eGFR – iGFR).

dUpdated Schwartz from Table 3.
eEquation III from Table 3: 39.1[ht/Scr]¹.8/CysC².29/BUN³.167[1.099ml/g]³/hd¹.4.188.

In our CKiD population, only 21% of the children were at Tanner stage IV or V, indicating that there was a very small proportion of fully developed adolescents in the study group. Examination of the estimating formulas in the children with Tanner stages IV through V failed to show substantial differences from the rest of the population (data not shown).
In addition, the method of cystatin C measurement in our study, using the Dako kit, may differ from the Siemens Dade-Behring determination, which has been reported as perhaps the most precise measurement of cystatin C available. The cystatin C–based estimate formula from Grubb et al.\textsuperscript{27} used the Dako kit as well, and this formula, regardless of whether adjusted for the CKiD training data set, did not perform well compared with our model III estimating equation.

With regard to other studies recently published using enzymatic creatinine method to estimate GFR via k[height/Scr] and measuring true GFR with a reliable method, k values in children and girls older than 13 were 0.500,\textsuperscript{29} 0.430,\textsuperscript{29} 0.470,\textsuperscript{30} and 0.460,\textsuperscript{31} compared with the 0.413 determined in this study. Note also that each of these other studies included individuals with normal and near-normal renal function, whereas children in CKiD were selected on the basis of reduced GFR with the mean and median approximating 40 ml/min per 1.73 m\textsuperscript{2}.

In sum, we have developed new formulas to estimate GFR in children with CKD. Precision and accuracy have been optimized by including in the equation endogenous Scr and cystatin C plus BUN, which explained 75.2% of the variability of iGFR, such that 87.7 and 45.5% of all eGFR values were within 30 and 10%, respectively, of simultaneously measured iGFR and complete data available on height, Scr, BUN, and cystatin C. Baseline sera and the serum separator tubes were shipped at room temperature to the Central Biochemistry Laboratory based in Rochester, NY, and sera for cystatin C were batched and sent quarterly to the nephrology laboratory at Children’s Mercy Hospital.

Studies and Assays
Before study blood was obtained for Scr, BUN, and cystatin C determinations, an aliquot was also obtained for HPLC determination of an iohexol blank. Scr (enzymatic) and BUN were analyzed centrally at the CKiD’s laboratory at the University of Rochester (G.J.S.) on an Advia 2400 (Siemens Diagnostics, Tarrytown, NY); cystatin C was measured centrally at the Children’s Mercy Hospital in Kansas (S. Hellerstein) by a turbidimetric assay (Cystatin C Kit K0071; DAKO SD, Copenhagen, Denmark). In a separate study, we showed that the Siemens Bayer Advia creatinine measurement closely agreed with an HPLC method traceable to reference isotope dilution mass spectrometry developed by the National Institute of Standards and Technology.\textsuperscript{34} The method of GFR determination using the plasma disappearance of iohexol in a two-compartment system has been previously reported.\textsuperscript{11} Iohexol (Omnipaque) was provided by GE Healthcare, Amersham Division (Princeton, NJ). No serious adverse events were noted in >700 studies.

Statistical Analysis
Nonparametric statistics (e.g., median and interquartile range) were used to describe the demographics of the study population and the components used in the calculation of iGFR. Regression analyses were performed in three stages. The first stage explored the univariate associations between BSA-unadjusted GFR and various markers of body size. It turned out that BSA and weight had the highest $R^2$ values and, more important, that BSA had a regression coefficient not different from 1. Hence, not only does the classical BSA-adjusted GFR provide a calibration to 1.73 m\textsuperscript{2}, but also its logarithm corresponds to the residuals of the regression of log(GFR) on log(BSA). After this stage, we used these residuals (the GFR of individuals of similar body size) as the outcome to determine the multivariate predictability of a number of markers of kidney function (Scr, height/Scr, cystatin C and BUN). In the final stage, we determined whether other variables, in-
including gender, race, Tanner stage, CKD cause, and body mass index, provided any additional predictive information. We explored potential interactions between variables when there were both significant main effects and biologically plausible interactions; for example, we explored the potential interaction between gender and height/Scr and between pubertal stage and Scr.

We used standard regression techniques for Gaussian data to determine the coefficients of the GFR estimating equations after logarithmic transformation of the continuous variables. All continuous independent variables were centered at the median values when entered into regression models. In this way, the models’ intercepts represent the expected value of GFR for the group of individuals with the constellation of predictors at the centering values (see the Results section).

The general regression model was of the form

\[
\log(\text{iGFR}[\text{ml/min}\text{per}1.73\text{m}^2]) = \log(a) + \log(X) + cZ + \epsilon,
\]

where \(X\) is a constellation of continuous predictors examined in stage 2 of analyses (e.g., height[m]/Scr[mg/dl], 1.8/cystatin C[mg/L], height[m]/1.4), \(Z\) is a constellation of categorical variables (e.g., gender, race) examined in stage 3 of analyses, and \(\epsilon\) follows a normal distribution with mean 0 and variance \(\sigma^2\) (where \(\sigma^2\) corresponds to the expected value of the mean square error); therefore, \(a\) represents the expected value of iGFR for the group whose values of the continuous predictors are at the median values of the study population (e.g., height[m]/Scr[mg/dl] = 1, cystatin C[mg/L] = 1.8, height[m] = 1.4) and whose categorical variables are at the reference categories (e.g., female).

The eGFR,

\[
eGFR = a[X]^2 \exp(c)^2,
\]

was obtained by using the expected values of the regression coefficients \((a, b,\) and \(c)\) along with specific values of the independent variables \((X\) and \(Z)\) for each individual. To assess the properties of the estimating equations, we calculated (1) the root mean square error, which measures the unexplained variability of iGFR; (2) the \(R^2\), which measures the percentage of the variability in iGFR explained by the predictors; (3) the correlation between the observed iGFR and eGFR; and (4) the percentage of eGFR values that were within 30 and 10% of the corresponding iGFR values calculated on the training sample (i.e., the one used to develop the equations) and on a testing sample composed of the 168 participants whose iGFR was measured 1 yr after the baseline visit. In addition, we tested the appropriateness of the Gaussian distribution for \(\log(\text{iGFR})\) against the rich family provided by members of the generalized gamma distribution.19

Comparison with Published Estimating Equations

We compared the estimating equations generated using the CKiD data with published Scr-based estimating equations by Counahan et al.2 and Leger et al.,25 with cystatin C–based equations by Filler et al.16 and Grubb et al.27 and with Scr– and cystatin C–based equations by Bouvet et al.28 and Zapitelli et al.21 using the originally published formulas and then by modifying the constants and coefficients in the published formulas after fitting them to the CKiD data. Applying each of these equations to the test group of 168 children, we determined the amount of bias from measured iGFR, 95% limits of agreement, correlation, and the percentage of estimates within 30 and 10% of measured iGFR.

ACKNOWLEDGMENTS

Data in this article were collected by the CKiD study with clinical coordinating centers (principal investigators) at Children’s Mercy Hospital and the University of Missouri–Kansas City (Bradley Warady, MD) and Johns Hopkins School of Medicine (Susan Furth, MD, PhD), and data coordinating center at the Johns Hopkins Bloomberg School of Public Health (Alvaro Muñoz, PhD) with the Central Biochemistry Laboratory at the University of Rochester (George J. Schwartz, MD). The CKiD is funded by the National Institute of Diabetes and Digestive and Kidney Diseases, with additional funding from the National Institute of Neurologic Disorders and Stroke, the National Institute of Child Health and Human Development, and the National Heart, Lung, and Blood Institute (U01-DK-66143, U01-DK-66174, and U01-DK-66116). The CKiD web site is located at http://www.statepi.jhsph.edu/ckid.

We are grateful to GE Healthcare, Amersham Division, for providing the CKiD study with iohexol (Omnipaque) for the GFR measurements. We are indebted to Paula Maier for coordinating the central laboratory and tracking the blood samples and to Brian Erway and Dr. Tai Kwong for skillfully developing and maintaining the iohexol HPLC assay at the University of Rochester Medical Center.

DISCLOSURES

None.

REFERENCES


