New Combined Serum Creatinine and Cystatin C Quadratic Formula for GFR Assessment in Children

Hassib Chehade,* François Cachat,* Anne-Sophie Jannot,† Blaise-Julien Meyrat,§ Dolores Mosig,* Daniel Bardy,§ Paloma Parvex,§ and Eric Girardin*

Summary
Background and objectives The estimated GFR (eGFR) is important in clinical practice. To find the best formula for eGFR, this study assessed the best model of correlation between sinistrin clearance (iGFR) and the solely or combined cystatin C (CysC)– and serum creatinine (SCreat)–derived models. It also evaluated the accuracy of the combined Schwartz formula across all GFR levels.

Design, setting, participants, & measurements Two hundred thirty-eight iGFRs performed between January 2012 and April 2013 for 238 children were analyzed. Regression techniques were used to fit the different equations used for eGFR (i.e., logarithmic, inverse, linear, and quadratic). The performance of each model was evaluated using the Cohen k correlation coefficient and the percentage reaching 30% accuracy was calculated.

Results The best model of correlation between iGFRs and CysC is linear; however, it presents a low k coefficient (0.24) and is far below the Kidney Disease Outcomes Quality Initiative targets to be validated, with only 84% of eGFRs reaching accuracy of 30%. SCreat and iGFRs showed the best correlation in a fitted quadratic model with a k coefficient of 0.53 and 93% accuracy. Adding CysC significantly (P<0.001) increased the k coefficient to 0.56 and the quadratic model accuracy to 97%. Therefore, a combined SCreat and CysC quadratic formula was derived and internally validated using the cross-validation technique. This quadratic formula significantly outperformed the combined Schwartz formula, which was biased for an iGFR=91 ml/min per 1.73 m².

Conclusions This study allowed deriving a new combined SCreat and CysC quadratic formula that could replace the combined Schwartz formula, which is accurate only for children with moderate chronic kidney disease. Clin J Am Soc Nephrol 9: 54–63, 2014. doi: 10.2215/CJN.00940113

Introduction
GFR is the best measure to assess kidney function, and the most accurate manner for its determination requires the use of endogenous markers (1–3) that are not easily available in clinical practice. Therefore, GFR is often estimated using endogenous markers, such as serum creatinine (SCreat) and serum cystatin C (CysC), with conflicting results regarding their performance (4–7).

This study assessed the best model of correlation between sinistrin clearance (iGFR), which is the gold standard method for determining measured GFR (mGFR) and CysC- or SCreat-derived formulas for determining estimated GFR (eGFR) in children. We first analyzed the best correlation model between iGFRs and CysC by evaluating the diagnostic accuracy of all previously published equation models based only on CysC for estimation of GFR. We then studied the best correlation between iGFRs and SCreat by evaluating the accuracy of the two recently published SCreat-based models by Schwartz et al. (8) and Gao et al. (9). We also analyzed the diagnostic accuracy of adding CysC and BUN to the best-fitted model of correlation. These led to deriving and validating the best combined SCreat and CysC accurate model/formula for eGFR in children. Finally, we compared the accuracy and validity of the combined SCreat and CysC Schwartz model and equation (10) to our combined SCreat and CysC fitted formula in a cohort of children with renal failure, as well as in children with normal renal function, about whom limited data are available in the literature.

Materials and Methods

Population
Two hundred forty-three iGFRs performed between January 2012 and April 2013 for 243 children were evaluated. All patients age 2–18.5 years referred to our laboratory unit for GFR measurement were included. Children with bladder dysfunction, those unable to void spontaneously, and those in whom bladder catheterization failed were excluded. Patients’ renal disorders and CKD classification are presented in Table 1. This study was approved by the local research ethics board and was conducted in accordance with the ethical standards of the Declaration of Helsinki.
and CysC
logarithmic Schwartz model and the combined SCreat

used to categorize iGFR and eGFR values. To compare

fi
compare the

(lowess function). Likelihood ratio tests were performed to

CysC was also assessed using a graphical representation

of observed values). Type of correlation between iGFR and

the percentage of estimated values within 10% and 30%

from both models. We also calculated the accuracy

(mean; median; first quartile, third quartile; minimum

and maximum) of RMSE values was reported. Additional

details on the methods and population are available in the

Supplemental Material.

Results

Two hundred forty-three patients were enrolled. In two

patients bladder catheterization failed, and three patients

had an allergic reaction and could not finish the sinistrin

clearance test. They were therefore excluded, which left

238 patients for final analysis. Patients’ demographic

characteristics are reported in Table 2. Ten children pre-

sented growth retardation for body weight (4.2% of all

Laboratory Analyses

SCreat was measured using the kinetic colorimetric

compensated Jaffe method, as reported by the manufacturer,

Roche Modular P system (Roche Diagnostics, Mannheim,

Germany), which was standardized to the isotope-dilution

mass spectrometry reference.

CysC was measured by particle-enhanced nephelometric

immunoassay on a BN ProSpec analyzer (Siemens Health-

care Diagnostics). The results were multiplied by 1.174 as

indicated in the Siemens customer bulletin to adjust to the

values obtained with the assay standardized to the new

traceable International Reference Preparation-ERM-

DA471/IFCC, as recommended by Kidney Disease Im-

prove Global Outcomes (11).

The IGFR (Inutest SPC, Fresenius Kabi Pharma, Austria)

was measured using the anthrone method by the manu-

facturer: Wright automatic method by Wright and Gann

(12), using an Autoanalyzer 3 system (high-resolution
digital colorimeter, SEAL Bran Luebbe, Norderstedt,

Germany).

Statistical Analyses

Statistical analyses were performed using R software,

version 2.15.2 (R Foundation for Statistical Computing,

Vienna, Austria). Linear regression techniques (function

"lm" in R software) were used to fit the different equations
to eGFR. We considered SCreat, CysC, age, and sex as

predictor variables for IGFR. In total, nine models adjusted

for age and sex with IGFR as the dependent variable and

CysC or SCreat as the independent variable were evalu-

ated. To assess the performance of the different models,

we calculated for each model the Cohen κ coefficient,

which measures agreement between two measurements;

values below and above 90 ml/min per 1.73 m² were

used to categorize iGFR and eGFR values. To compare κ

coefficients between two models, we used a permutation

procedure by randomly permuting the estimated values

from both models. We also calculated the accuracy (i.e.,

the percentage of estimated values within 10% and 30% of

observed values). Type of correlation between IGFR and

CysC was also assessed using a graphical representation

(lowess function). Likelihood ratio tests were performed to

compare the fit of the combined SCreat and CysC–based

logarithmic Schwartz model and the combined SCreat

and CysC–based quadratic model. The cutoff of the

applicability of the new combined Schwartz formula was
determined by applying the circular binary segmentation

method (9,13,14) to formula residuals. The circular binary

segmentation method allows segmenting of data through

change-point detection using a likelihood ratio statistic,

which tests the null hypothesis that there is no change,

against the alternative hypothesis that there is exactly

one change at each location.

To test for normality of model residuals, we used two

approaches: The first one involved the D’Agostino and

Pearson omnibus normality test (package fBasics, R soft-

ware), and the second one graphically assessed the

normality assumption by plotting residuals on a Q-Q plot.
The Q-Q plot shows the quantile of the distribution of

the new quadratic formula residuals (y-axis) versus the

quantile of a Gaussian distribution (x-axis) with a mean

of 0 and an SD of 1.

To check for internal validity of our models estimates, we

performed a cross-validation technique called repeated

random subsampling validation, also known as Monte

Carlo cross-validation (15). This involves randomly divid-
ing the data into two samples: the training set (two thirds

of the sample), on which the model was fitted, and the

testing set (one third of the sample), on which the model

was evaluated. The cross-validation was done with 1000

bootstrap replications. For each replication, the model was

fit to the training data; the root mean square error (RMSE)

was calculated using this fitted model for the training set

and then for the testing set. We therefore developed for
each replication our estimating equation from the training

data and validated it on the validation data. The distribu-
tion (mean; median; first quartile, third quartile; minimum

and maximum) of RMSE values was reported. Additional

table 1. Patients’ renal disorders and CKD classification

<table>
<thead>
<tr>
<th>Cause of CKD</th>
<th>CKD Stage I</th>
<th>CKD Stage II</th>
<th>CKD Stage III</th>
<th>CKD Stage IV and V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obstructive or reflux uropathy</td>
<td>66</td>
<td>56</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Congenital and acquired single kidney</td>
<td>20</td>
<td>25</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Polycystic kidney disease</td>
<td>7</td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Glomerulopathies</td>
<td>6</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Hemolytic-uremic syndrome</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Metabolic disease</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Post-chemotherapy</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>8</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Total (n=238)</td>
<td>106</td>
<td>105</td>
<td>23</td>
<td>4</td>
</tr>
</tbody>
</table>

Values are the numbers of patients. CKD stages I, II, III, IV, and V denote GFR≥90, 60–89, 30–59, 15–29, and <15 ml/min per 1.73 m², respectively.
patients), 17 presented growth retardation for height (7.1% of all patients), and 20 presented combined growth retardation for both body weight and height (8.4% of all patients).

Correlation between iGFR and CysC

The correlation between iGFR and CysC was analyzed in the whole population using logarithmic, inverse, quadratic, and simple linear fitted models, as follows: log[iGFR] = a + b log(CysC) + a log(age) + b (sex), iGFR = a + b(1/CysC) + a (age) + b (sex), iGFR = a + b(CysC) + γ(CysC)² + a (age) + b (sex), and iGFR = a + b(CysC) + a (age) + b (sex), respectively. Regression coefficients were calculated independently of the published coefficient equations (5, 6, 16–19). As shown in Figure 1, the scatter plot that graphically appears to be the closest to a straight line was the one performed without transformation of CysC, in favor of a linear relationship between CysC and iGFR. Furthermore, κ coefficient for the linear model was higher than for the inverse and logarithmic models (Table 3). However, accuracies of all these different models were poor: We found that 42%, 42%, 42%, and 41% of eGFRs were within 10% of iGFR and 84%, 83%, 84%, and 84% of eGFRs were within 30% of iGFR for the linear, inverse, logarithmic, and quadratic correlations, respectively.

Correlation between iGFR and SCreat

Schwartz et al. proposed a linear formula for eGFR using SCreat (8) as follows: eGFR = 0.413 × height (Ht)/SCreat. We first analyzed, in the whole population, the correlation between iGFR and this linear model independent of the defined formula’s constant (0.413) as follow: iGFR = a + b(Ht/SCreat) + a (age) + b (sex). This linear model achieved accuracies within 10% and 30% of the observed values of 49% and 90%, respectively, compared with 45% and 88%, respectively, with the 0.413 Schwartz coefficient formula. Second, we fit the relationship between iGFR and the variable ratio of Ht/SCreat in a quadratic model, as proposed by Gao et al. (9), as follows: iGFR = a + b(Ht/SCreat) + γ(Ht/SCreat)² + a (age) + b (sex). We obtained a better model performance than the one obtained from the linear model, with 93% of eGFR values being within 30% of iGFR values and a highly significant (P<0.001) likelihood ratio test between these two models (Table 3). The κ coefficient was also increased for the quadratic model compared with the linear one (0.53 versus 0.46). Therefore, adding a quadratic term to the linear SCreat-based model increased the fit of the linear model.

Correlation between iGFR and New Quadratic SCreat and CysC–Based Model with CysC

With use of the whole population, adding CysC to the SCreat-based quadratic model significantly increased model fit with a significant (P<0.001) likelihood ratio test. The analysis of the accuracy of this new quadratic SCreat and CysC–based model (Table 3) showed that for an accuracy of 10% and 30%, 53% and 97% of eGFRs were accurate, respectively. This combined SCreat and CysC quadratic model also achieved a better κ coefficient—0.56—than did the solely SCreat quadratic model.

Correlation between iGFR and New Quadratic SCreat and CysC–Based Model with BUN

With use of the whole population, adding BUN to the combined quadratic SCreat and CysC–based model did not increase model fit, with a nonsignificant (P=0.23) log-likelihood ratio chi-square difference between the two models with and without BUN. Table 3 shows that the performances of the models with or without BUN were equal. Therefore, the BUN variable was not added to the final chosen model.

Quadratic SCreat and CysC–Based Model Validation and New Combined Quadratic Formula: Internal Model Validation

A Monte Carlo cross-validation technique (15) as described in the statistical analyses section was performed. The cross-validation was done with 1000 bootstrap replications. In each replication, the RMSE was calculated using the fitted quadratic model for the training set and for the testing set. This allows comparing the RMSE between the training set and testing set (Table 4). The mean of RMSE obtained for the testing set was close to the one obtained for the training set: 12.8 for the new combined quadratic model in the testing set versus 12.04 for the combined quadratic model in the training set. This showed good internal consistency of our estimates.
Figure 1. | Graphs depicting sinistrin clearances (iGFRs) according to cystatin C (CysC) (A), 1/CysC (B), logarithm of CysC (C), and square of CysC (D), using a locally weighted polynomial regression. The solid line represents a local linear polynomial fit (function lowess, package stats).

Table 3. Analysis of estimated GFR correlations with sinistrin clearance and accuracy of studied models

<table>
<thead>
<tr>
<th>Fitted Models and Schwartz Formula</th>
<th>Cohen ( \kappa ) Coefficient</th>
<th>Accuracy within 10% of iGFR (%)</th>
<th>Accuracy within 30% of iGFR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear CysC-based model</td>
<td>0.24</td>
<td>42</td>
<td>84</td>
</tr>
<tr>
<td>Inverse CysC-based model</td>
<td>0.21</td>
<td>42</td>
<td>83</td>
</tr>
<tr>
<td>Logarithmic CysC-based model</td>
<td>0.23</td>
<td>42</td>
<td>84</td>
</tr>
<tr>
<td>Quadratic CysC-based model</td>
<td>0.24</td>
<td>41</td>
<td>84</td>
</tr>
<tr>
<td>Linear SCreat-based model</td>
<td>0.46</td>
<td>49</td>
<td>90</td>
</tr>
<tr>
<td>Quadratic SCreat-based model</td>
<td>0.53</td>
<td>52</td>
<td>93</td>
</tr>
<tr>
<td>New combined SCreat and CysC–based quadratic model</td>
<td>0.56</td>
<td>53</td>
<td>97</td>
</tr>
<tr>
<td>Combined SCreat, CysC, and BUN–based quadratic model</td>
<td>0.56</td>
<td>53</td>
<td>97</td>
</tr>
<tr>
<td>Combined SCreat and CysC–based logarithmic Schwartz model</td>
<td>0.49</td>
<td>54</td>
<td>96</td>
</tr>
<tr>
<td>Combined SCreat and CysC Schwartz formula</td>
<td>0.15</td>
<td>35</td>
<td>86</td>
</tr>
</tbody>
</table>

iGFR, sinistrin clearance; SCreat, serum creatinine; CysC, cystatin C.
Table 4. Root mean square error distribution in bootstrap samples using new combined serum creatinine and cystatin C quadratic model (232 degrees of freedom) and combined serum creatinine and cystatin C–based logarithmic Schwartz model (231 degrees of freedom) for both training and testing sets

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>10.37</td>
<td>10.73</td>
<td>-0.83</td>
<td>9.76</td>
<td>09.93</td>
<td>-2.11</td>
</tr>
<tr>
<td>First quartile</td>
<td>11.70</td>
<td>12.21</td>
<td>3.23</td>
<td>12.05</td>
<td>12.46</td>
<td>1.94</td>
</tr>
<tr>
<td>Median</td>
<td>12.08</td>
<td>12.64</td>
<td>4.46</td>
<td>12.79</td>
<td>13.33</td>
<td>4.04</td>
</tr>
<tr>
<td>Mean</td>
<td>12.04</td>
<td>12.61</td>
<td>4.35</td>
<td>12.80</td>
<td>13.39</td>
<td>4.89</td>
</tr>
<tr>
<td>Third quartile</td>
<td>12.43</td>
<td>13.06</td>
<td>5.63</td>
<td>13.56</td>
<td>14.36</td>
<td>7.10</td>
</tr>
</tbody>
</table>

Relative reduction of RMSE was calculated by subtracting for each replicate the RMSE obtained from the combined logarithmic model and the one obtained from the combined quadratic model. RMSE, root mean square error.

On the basis of results described above, we derived from the whole population a new quadratic SCreat and CysC–based formula for eGFR. Estimated coefficients and 95% confidence intervals of this new formula are reported in Table 5. The new quadratic formula for female and male patients is as follows:

Female patients:

\[
0.42 \times (Ht/SCreat) - 0.04 \times (Ht/SCreat)^2 - 14.5 \times CysC + 0.69 \times age + 18.25
\]

Males:

\[
0.42 \times (Ht/SCreat) - 0.04 \times (Ht/SCreat)^2 - 14.5 \times CysC + 0.69 \times age + 21.88
\]

(where Ht is in cm, SCreat in mg/dl, CysC in mg/L, and age in years).

Table 5. Estimated coefficients with 95% confidence intervals of the new combined serum creatinine and cystatin C quadratic model

<table>
<thead>
<tr>
<th>New Combined SCreat and CysC Quadratic Model</th>
<th>Coefficient Estimate (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>18.25 (-4.84 to 41.34)</td>
</tr>
<tr>
<td>Ht/SCreat</td>
<td>0.42 (0.29 to 0.56)</td>
</tr>
<tr>
<td>((Ht/SCreat)^2)</td>
<td>-0.04 (-0.07 to -0.02)</td>
</tr>
<tr>
<td>CysC</td>
<td>-14.52 (-21.43 to -7.61)</td>
</tr>
<tr>
<td>Age</td>
<td>0.69 (0.28 to 1.1)</td>
</tr>
<tr>
<td>Sex</td>
<td>3.63 (0.55 to 6.72)</td>
</tr>
</tbody>
</table>

SCreat, serum creatinine expressed in mg/dl; CysC, cystatin C expressed in mg/L; Ht, height expressed in cm.

While normal distribution for residuals was rejected according to the D’Agostino and Pearson omnibus normality test, the residual Q-Q plot did not show a strong departure from normality assumption (Figure 2). Of note, no observation achieved a Cook distance >4/n (n is the number of observations), which is classically (20) admitted as the cutoff value to use for spotting highly influential points.

Is the Combined SCreat and CysC Schwartz Equation (2012) Valid in Our Study Sample?

We fit the combined SCreat, CysC, and BUN–based logarithmic model proposed by Schwartz et al. in 2012, which includes six variables (height, age, sex, and three endogenous biomarkers: SCreat, CysC, and BUN) as follow: log(iGFR) = α + βlog(Ht/SCreat) + γlog(Ht) + δlog(CysC) + θlog(BUN) + alog(age) + b(sex). We observed that in the whole population, the combined logarithmic Schwartz model had a significantly lower κ coefficient than did the new combined SCreat and CysC quadratic model (0.49 versus 0.56; P=0.01 using a permutation procedure). Both combined models had very similar accuracies; 96% and 97% of eGFRs were within 30% of iGFR, respectively. With use of estimated coefficients proposed by Schwartz et al. (10) for the entire population, a much lower κ coefficient was obtained (0.15); for an accuracy of 30%, only 86% of eGFRs reached this accuracy level (Table 3). Meanwhile, when we restricted our population to individuals with iGFR values ranging from 15 to 75 ml/min per 1.73 m², accuracy increased for the Schwartz formula; 90% of estimated values were within 30% of the iGFRs.

We also compared internal consistency of the Schwartz combined logarithm model to that of the quadratic combined model (Table 4) using the training and the testing datasets. The lowest RMSE, and therefore the best fit, was obtained using the new combined SCreat and CysC–based quadratic model. Mean RMSE between the training and the testing set for the logarithmic model...
increased at the same percentage as for the new combined quadratic model (approximately 6%), meaning that the logarithmic model suggested by Schwartz (without the defined Schwartz coefficients) was not biased in our study population.

To define the cutoff of applicability of the combined Schwartz formula, we applied the circular binary segmentation method for residual values (iGFR - eGFR). This method demonstrates the presence of three points of change corresponding to iGFRs of 75, 91, and 126 ml/min per 1.73 m²; as a result, four segments were identified (Figure 3). When we applied the circular binary segmentation method to residual values from the new combined quadratic formula, we also detected (Figure 4) three points of change corresponding to iGFRs of 81, 96, and 126 ml/min per 1.73 m². In consequence, four segments were identified. Detailed results for each segment regarding the mean difference between iGFRs and eGFRs and the accuracy within 30% of iGFRs are reported in Table 6. We observed that mean absolute value of the segment (i.e., the mean bias) was greater for the combined Schwartz formula than for the combined quadratic formula and that accuracy was also better for the latter formula in all segments except for iGFRs < 91 ml/min per 1.73 m², corresponding to the population from which the combined Schwartz formula was derived. In addition, results showed that the combined Schwartz formula was biased for iGFR ≥ 91 ml/min per 1.73 m². For patients with iGFRs above that precise cutoff, the mean difference between iGFRs and eGFRs was significantly wider and the accuracy within 30% of iGFRs was significantly lower compared with patients with iGFRs < 91 ml/min per 1.73 m².

Is There a Practical Reason for Clinicians to Use the New Combined Quadratic Formula?

From a practical point of view, it is mostly important for clinicians to differentiate children with normal (i.e., ≥ 90 ml/min per 1.73 m²) from those with abnormal GFR. We divided our cohort into two categories: below and above 90 ml/min per 1.73 m². Performances of the all-fitted models and of the combined Schwartz formula were evaluated by calculating the sensitivity, specificity, and area under the curve (AUC) (Table 7). Using a permutation procedure, we also calculated the difference in the κ concordance coefficient between the new combined quadratic model/formula and the SCreat-based Schwartz model. Results show that although the AUC of the linear SCreat-based model is similar to that of the new combined quadratic model and formula, its specificity of 60% is much lower than the one obtained with the new combined quadratic model or formula; as a result it could lead to misclassification of children with normal GFRs into abnormal GFRs categories. In addition, the κ coefficient obtained from the new combined quadratic model or formula was significantly better than the one obtained from the linear SCreat-based model (P < 0.001). Therefore, in clinical practice, the new combined quadratic formula achieves the best agreement with iGFR.

To facilitate the reader’s understanding and interpretation of our findings, Table 8 summarizes the main performances of the SCreat-based Schwartz equation (8), the SCreat-based quadratic formula proposed by Gao et al. (9), the combined logarithmic Schwartz model with coefficients fitted to our population, the combined logarithmic Schwartz formula with its original coefficients (10), and the new combined quadratic formula.

![Q-Q plot for residuals of the combined serum creatinine (SCreat) and cystatin C (CysC) quadratic model.](image-url)
Figure 3. | Residuals versus sinistrin clearances (iGFRs) and corresponding segments by circular binary segmentation method using the combined serum creatinine (SCreat) and cystatin C (CysC)–based Schwartz formula. eGFR, estimated GFR.

Figure 4. | Residuals versus sinistrin clearances (iGFRs) and corresponding segments by circular binary segmentation method using the new combined serum creatinine (SCreat) and cystatin C (CysC)–based quadratic formula. eGFR, estimated GFR.
Several formulas for eGFR in children were recently developed on the basis of SCreat (8,9), CysC (5,6,16–19), or combined SCreat and CysC measurements (10). They all show limitations in accuracy. We found that the best-fit model of correlation between iGFR and CysC is linear, but it is still far below the KDOQI targets to be validated, with $<90\%$ of estimated GFRs reaching $30\%$ accuracy (21).

### Table 6. Mean difference between sinistrin clearances and estimated GFRs and accuracy within $30\%$ of sinistrin clearance for each segment obtained by applying circular binary segmentation for all bias values (sinistrin clearance – estimated GFR) using combined Schwartz and new combined quadratic formulas

<table>
<thead>
<tr>
<th>CBS Segments and Corresponding iGFRs (ml/min per 1.73 m$^2$)</th>
<th>Mean Difference between iGFRs and eGFRs (ml/min per 1.73 m$^2$)</th>
<th>Accuracy within $30%$ of iGFR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using combined Schwartz formula</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First segment: iGFRs&lt;75</td>
<td>0</td>
<td>90</td>
</tr>
<tr>
<td>Second segment: 75≤iGFRs&lt;91</td>
<td>9</td>
<td>95</td>
</tr>
<tr>
<td>Third segment: 91≤iGFRs&lt;126</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>Fourth segment: iGFRs≥126</td>
<td>NE</td>
<td>NE</td>
</tr>
<tr>
<td>Using new combined quadratic formula</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First segment: iGFRs&lt;81</td>
<td>−7</td>
<td>93</td>
</tr>
<tr>
<td>Second segment: 81≤iGFRs&lt;96</td>
<td>−1</td>
<td>100</td>
</tr>
<tr>
<td>Third segment: 96≤iGFRs&lt;126</td>
<td>9</td>
<td>100</td>
</tr>
<tr>
<td>Fourth segment: iGFRs≥126</td>
<td>NE</td>
<td>NE</td>
</tr>
</tbody>
</table>

CBS, circular binary segmentation; iGFR, sinistrin clearance; eGFR, estimated GFR; NE, not evaluated (only four observations identified in that CBS segment, precluding any statistical analysis).

### Table 7. Sensitivity, specificity, and area under curve of all-fitted models and combined quadratic and Schwartz formulas to differentiate children with normal ($≥90$ ml/min per 1.73 m$^2$) from those with abnormal GFR

<table>
<thead>
<tr>
<th>Fitted Models and Schwartz Formula</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>AUC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear CysC-based model</td>
<td>63</td>
<td>61</td>
<td>72</td>
</tr>
<tr>
<td>Inverse CysC-based model</td>
<td>68</td>
<td>53</td>
<td>71</td>
</tr>
<tr>
<td>Logarithmic CysC-based model</td>
<td>70</td>
<td>54</td>
<td>71</td>
</tr>
<tr>
<td>Quadratic CysC-based model</td>
<td>63</td>
<td>62</td>
<td>72</td>
</tr>
<tr>
<td>Linear SCreat-based model</td>
<td>85</td>
<td>60</td>
<td>86</td>
</tr>
<tr>
<td>Quadratic SCreat-based model</td>
<td>80</td>
<td>74</td>
<td>86</td>
</tr>
<tr>
<td>New combined SCreat and CysC–based quadratic model</td>
<td>80</td>
<td>77</td>
<td>86</td>
</tr>
<tr>
<td>Combined SCreat, CysC, and BUN–based quadratic model</td>
<td>80</td>
<td>77</td>
<td>86</td>
</tr>
<tr>
<td>New combined SCreat and CysC–based quadratic formula</td>
<td>80</td>
<td>77</td>
<td>86</td>
</tr>
<tr>
<td>Combined SCreat and CysC–based logarithmic Schwartz model</td>
<td>81</td>
<td>68</td>
<td>86</td>
</tr>
<tr>
<td>Combined SCreat and CysC Schwartz formula</td>
<td>94</td>
<td>20</td>
<td>78</td>
</tr>
</tbody>
</table>

SCreat, serum creatinine; CysC, cystatin C.

### Table 8. Main performance of serum creatinine–based Schwartz formula, serum creatinine–based quadratic formula proposed by Gao et al., combined logarithmic Schwartz model with coefficients fitted to current population, combined logarithmic Schwartz formula with original coefficients, and new combined quadratic formula

<table>
<thead>
<tr>
<th>Variable</th>
<th>Accuracy within 10% of iGFR (%)</th>
<th>Accuracy within 30% of iGFR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCreat-based Schwartz formula</td>
<td>45</td>
<td>88</td>
</tr>
<tr>
<td>SCreat-based quadratic formula</td>
<td>44</td>
<td>90</td>
</tr>
<tr>
<td>Combined SCreat and CysC–based logarithmic Schwartz model</td>
<td>54</td>
<td>96</td>
</tr>
<tr>
<td>Combined SCreat and CysC Schwartz formula</td>
<td>35</td>
<td>86</td>
</tr>
<tr>
<td>New combined SCreat and CysC–based quadratic formula</td>
<td>53</td>
<td>97</td>
</tr>
</tbody>
</table>

iGFR, sinistrin clearance; SCreat, serum creatinine; CysC, cystatin C.

**Discussion**

Several formulas for eGFR in children were recently developed on the basis of SCreat (8,9), CysC (5,6,16–19), or combined SCreat and CysC measurements (10). They all show limitations in accuracy. We found that the best-fit model of correlation between iGFR and CysC is linear, but it is still far below the KDOQI targets to be validated, with $<90\%$ of estimated GFRs reaching $30\%$ accuracy (21). Regarding the solely SCreat-based formulas for eGFR, the linear model proposed by Schwartz et al. is inaccurate...
This study has some limitations. First, very few patients had CKD stage IV and V. However, the circular binary segmentation method applied for all GFR values using the new combined quadratic formula did not detect any point of change in patients with CKD stage III or higher. Additional studies should confirm the validity of the new combined quadratic formula in children with CKD stages IV and V. Second, we found that the best correlation between iGFR and CysC was linear. However, few patients in our study had a CysC > 2.5 mg/L; therefore, we cannot exclude the presence of a one-phase exponential decay or a hyperbola that our data did not demonstrate. Third, Schwartz et al. (10) found an increased bias in the eGFR for heavier patients. The small number of overweight and obese children or children with growth retardation prevented any meaningful evaluation of the new quadratic formula in this selective population. Fourth, the new combined quadratic formula does not overcome limitations of SCreat measurements in some patients, which depend on muscle mass levels. It is important to note that muscle mass was not evaluated in our population, but no infant was diagnosed with myopathy or malnutrition.

Finally, we cannot rule out that a decrease in the performance of the solely SCreat or the combined Schwartz formulas could occur secondary to the different SCreat and/or CysC measurement methods. However, in our study, SCreat measurement was performed using the compensated Jaffe technique standardized against isotope-dilution mass spectrometry method, with inter- and intra-assay coefficients of variation far below the current recommendations of the Laboratory Working Group of the National Kidney Disease Education Program (25). We also compared the compensated Jaffe technique to the enzymatic technique used by Schwartz et al. and found an average difference between the enzymatic and the Jaffe methods of 0.99 μmol/L (95% confidence interval, −6.586 to 8.566 μmol/L), which shows that both methods are closely aligned. In our opinion, the differences in SCreat measurements must have a small error that should minimally affect the performance of the SCreat-based Schwartz formulas in our population. Regarding CysC measurement, and contrary to our standardized method of analysis, Schwartz et al. in their cohort could not use the standardized CysC calibrators, which could decrease the performance of their formula in a population such as ours.

Acknowledgments
The authors would like to thank all colleagues at the Lausanne University Hospital who graciously contributed to this study. We are grateful for Mr. Christian Lehmann, Mr. Joel Stauber, Mr. Ziad Daher, Mr. Manuel Petter, and Mrs. Valerie Blanc for developing the combined quadratic formula website.

Disclosures
None.

References


Received: February 15, 2013 Accepted: September 1, 2013

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

This article contains supplemental material online at http://cjasn.asnjournals.org/lookup/suppl/doi:10.2215/CJN.00940113/-/DCSupplemental.
Correction

Because of a publisher error, in the original manuscript entitled, “New Combined Serum Creatinine and Cystatin C Quadratic Formula for GFR Assessment in Children,” which appeared in the January 2014 issue of CJASN, the formulas were printed incorrectly on page 58 and in Table 5 as boldly below. CJASN apologizes to the authors for printing these errors.

Page 58:

Female patients:
0.42 × (Ht/SCreat) − 0.04 × (Ht/SCreat)^2 − 14.5 × CysC + 0.69 × age + 18.25
Males:
0.42 × (Ht/SCreat) − 0.04 × (Ht/SCreat)^2 − 14.5 × CysC + 0.69 × age + 21.88.

Table 5:

−0.04 (−0.07 to −0.02).

The correct formulas are listed below.

Page 58:

The new Combined Quadratic formula for female and male patients is as follows:

Female patients:
0.42 × (Ht/SCreat) − 0.0004 × (Ht/SCreat)^2 − 14.5 × CysC + 0.69 × age + 18.25
Males:
0.42 × (Ht/SCreat) − 0.0004 × (Ht/SCreat)^2 − 14.5 × CysC + 0.69 × age + 21.88,

where height (Ht) is in centimeters, serum creatinine (SCreat) is in milligrams per deciliter, cystatin C (CysC) is in milligrams per liter, and age is in years.

Table 5:

−0.0004 (−0.0007 to −0.0002).

Published online ahead of print. Publication date available at www.cjasn.org.
New combined serum creatinine and cystatin C Quadratic formula for glomerular filtration rate assessment in children

Chehade H., Cachat F., Jannot AS., Meyrat BJ., Mosig D., Bardy D., Parvex P., Girardin E.

Lausanne University Hospital, Department of Pediatrics, Division of Pediatric Nephrology, CRC & Division of clinical-epidemiology, Department of health and community medicine University of Geneva & University Hospitals of Geneva. Lausanne University Hospital, Department of Pediatrics, Division of Pediatric Surgery. Lausanne University Hospital, Central Chemistry Laboratory. Geneva University Hospital, Department of Pediatrics, Division of Pediatric Nephrology

Materials and methods

Population

This study has been approved by the local research ethics board. Two hundred and forty-three sinistrin clearances (iGFRs) performed between January 2012 and April 2013 for 243 children were evaluated. Indications for clearance performance were left at the discretion of the referring physician (pediatric nephrologists or urologists). All patients aged between 2 and 18.5 years old, referred to our laboratory unit for glomerular filtration rate (GFR) measurement were included. Children with bladder dysfunction, unable to void spontaneously, and in whom bladder catheterization failed, were excluded. Proper emptying of the bladder was also evaluated by comparing the urine output with its osmolality. A decreasing diuresis with a concomitant decreasing urine osmolality was an indication of poor bladder emptying, and the child was then excluded from the study for technical purpose. Our study did not include transplant patients, already enrolled in another study. Body weight (BW) and height (Ht) were expressed in absolute values and in percentiles (P) according to Swiss pediatric growth charts (1). Body mass index (BMI) was calculated as a person’s weight in kilograms divided by the square of his height in meters (kg/m²). Growth retardation was
defined by a Ht and/or a BW below P10. Overweight was defined by a BMI between 25 and 29.9 kg/m² and obesity was defined by a BMI of 30 kg/m² or higher. Patients’ renal disorders and chronic kidney disease (CKD) classification are presented in table 1.

Analytical analysis

Serum creatinine (SCreat) calibration:

SCreat was measured using the kinetic colorimetric compensated Jaffe method as reported by the manufacturer Roche Modular P system (Roche Diagnostics, Mannheim, Germany) that was standardized to the isotope-dilution mass spectrometry (IDMS) reference. The method was calibrated with the calibrator and procedures were described by the Roche Diagnostics. The IDMS traceability involves two points calibration (target values at 0 µmol/l and 360-390 µmol/l depending on the calibrator lot) and the subtraction of 26 µmol/l from the results in order to compensate for the non specific chromogens. The inter-assay coefficients of variation (CVs) obtained in the laboratory with the internal quality controls were 3.9% at 45.7 µmol/l and 2.4% at 108 µmol/l. The intra-assay CVs were 3.3% at 44.5 µmol/l and 0.7% at 148 µmol/l. The laboratory is participating in an external quality assessment scheme.

Cystatin C (CysC) calibration:

CysC was measured by particle-enhanced nephelometric immunoassay (PENIA) on BN ProSpec analyser (Siemens Healthcare Diagnostics). The results were multiplied by 1.174 as indicated in the Siemens customer bulletin to adjust to the values obtained with the assay standardized to the new traceable International Reference Preparation (IRP)-ERM®-DA471/IFCC as recommended by the KDIGO (Kidney Disease Improve Global Outcomes) (2). Normal reference range of CysC was 0.55 to 1.06 mg/l. The inter-assay CVs were 4.95% at 1.13 mg/l and 3.30% at 4.91 mg/l. The intra-assay CVs were 2.0% at 1.71 mg/l and 2.3% at 5.37 mg/l. Total measurement imprecision was 0.097 mg/l at 1.14 mg/l.
**GFR measurement:**

All patients fasted for an overnight before the day of investigation, and drugs interfering with the sinistrin measurement were omitted before and during the test. Sinistrin clearance (iGFR) was obtained as follows: Two intravenous catheters were inserted on admission, one in each arm and a loading dose of sinistrin 25% was administrated according to the (Inutest SPC - Fresenius Kabi Pharma Austria GmbH) protocol. The loading dose was calculated in order to obtain a required plasma concentration of 200-250 mg/l, as follow: Loading sinistrin dose = Required plasma concentration x Estimated sinistrin distribution volume. The estimated sinistrin distribution volume corresponds to the extracellular volume and amounts to 15% of the body weight. Subsequently, sinistrin was continuously infused over 90 minutes at a rate given by the required inulin plasma concentration (200-250 mg/l) and the estimated GFR as follow: Infusion rate in mg/mn = (Required plasma concentration / 1000 ) x Estimated GFR in ml/mn. Water diuresis was induced by oral administration of 20 ml/kg of water (maximum 1200ml) in the first hour followed by 3 ml/kg/h of water. This was combined with an intravenous infusion of 0.9% sodium chloride (maximum 300 ml) every 30 minutes. After a 90-minutes equilibration period, 3 timed-urine samples were collected every 30 minutes, according to the manufacturer’s protocol (Inutest SPC - Fresenius Kabi Pharma Austria GmbH), with a blood test in the middle of each urine collection. Sinistrin was measured using the anthrone method by the manufacturer Wright’s automatic Wright and Gann (3), using an Autoanalyzer 3 system (High resolution digital colorimeter of SEAL; BRAN+LUEBBE, Norderstedt, Germany). The method was calibrated with five points calibrations (targets values at 10 mg/100 ml, 20mg/100ml, 30 mg/100ml, 40 mg/100ml, 50 mg/100ml) with coefficients correlations of 0.9993±0.0005. The procedure is automatized with software (Bran+ Luebbe, AACE 6.03) which automatically includes corrections for baseline, carryover, sensitivity drift and dilution factor. For the serum, the intra-assay coefficients of
variation (CVs) obtained in our laboratory with the internal quality controls were 2.44% at 10mg/100ml, 1.47% at 30mg/100ml and 0.94% at 40 mg/100ml, while for urine the intra-assay CVs were 1.71% at 10mg/100ml, 1.22% at 30mg/100ml and 1.07% at 50 mg/100ml. The inter-assay CVs obtained in the internal laboratory standards were 2.35% at 10mg/100ml, 2.23% at 30mg/100ml and 0.87% at 50 mg/100ml.

iGFR was calculated as the mean of the three clearance periods. When sinistrin clearance difference between 2 periods exceeded 20%, that period was excluded, and iGFR was calculated as the mean of the 2 valid periods.

**Statistical analysis**

Statistical analysis was performed using R software, version 2.15.2 (R Foundation for Statistical Computing, Vienna, Austria). Population demographics were summarized with median (interquartile range), minimal and maximal values for continuous characteristics and percentages for categorical characteristics. Linear regression techniques (function “lm” in R software) were used to fit the different equations to estimated GFR (eGFR). We considered as predictors parameters for iGFR, the SCreat, the CysC, the age and the sex. In total, 9 models adjusted for age and sex with iGFR as dependant variable and CysC or SCreat as independent variable were evaluated. To assess the performance of the different models, we calculated for each model the Cohen’s Kappa (κ) coefficient that measures agreement between two measurements, using values below and above 90 ml/mn per 1.73m² to categorize iGFR and eGFR values. To compare κ coefficient between two models, we used a permutation procedure by randomly permuting the estimated values from both models. We also calculated the accuracy, i.e. the % of estimated values within 10% and 30% of observed values. Type of correlation between iGFR and CysC was also assessed using a graphical representation (function lowess). Likelihood ratio tests were performed to compare the fit of the combined SCreat and CysC based logarithmic Schwartz model and the combined SCreat and CysC
based quadratic model. Significance was defined as $p \leq 0.05$. The cut-off of the applicability of the new combined Schwartz formula was determined by applying the circular binary segmentation (CBS) method (4, 5, 6) to formula residuals, as follows: If we consider $R_1, R_2, \ldots, R_n$ the residuals (iGFR-eGFR) of iGFR with iGFR sorted in ascending order and let $S_i = R_1 + R_2 + \ldots + R_i$, $1 \leq i \leq n$, be the partial sums, the likelihood ratio statistic for testing the null hypothesis that there is no change against the alternative that there is exactly one change at unknown location $I$ is given by $Z_B = \max_{1 \leq i < n} |Z_i|$, where $Z_i = \{1/i + 1/(n-i)\} - 0.5 \{S_i / I - (S_n - S_i)/(n-i)\}$. The null hypothesis of no change is rejected if the statistic exceeds the upper $\alpha$th quantile of the null distribution of $Z_B$ and the location of the change-point is estimated to be $i$ such that $Z_B = |Z_i|$. The CBS method allows to segment data through change-point detection using a maximal t-test.

To test for normality of model residuals, we used two approaches: the first one was to perform D’Agostino and Pearson omnibus normality test (package fBasics, R software) and the second one to graphically assess the normality assumption by plotting residuals on a Q-Q plot. The Q-Q plot shows the quantile of the distribution of the new quadratic formula residuals (y-axis) versus the quantile of a Gaussian distribution (x-axis) with mean 0 and standard deviation 1.

To check for internal validity of our models estimates, we performed a Cross-validation technique called repeated random sub-sampling validation, also known as Monte Carlo Cross-validation (7). This involves dividing randomly the data into two samples: the training set (2/3 of the sample), on which the model was fitted, and the testing set (1/3 of the sample), on which the model was evaluated. The cross-validation was done with 1000 bootstrap replications. For each replication, the model was fitted to the training data, and the root mean square error (RMSE) was calculated using this fitted model for the training set, and then for the testing set. We therefore developed for each replication our estimating equation from the
training data and validated it on the validation data. The distribution (mean, median, first quartile, third quartile, minimum and maximum) of RMSE values was reported.

References


7. Witten IH, Frank E: Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2005
### Table 1:

<table>
<thead>
<tr>
<th>Etiologies:</th>
<th>CKD stage I</th>
<th>CKD stage II</th>
<th>CKD stage III</th>
<th>CKD stage IV and V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obstructive or reflux uropathy</td>
<td>66</td>
<td>56</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Congenital and acquired single kidney</td>
<td>20</td>
<td>25</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Polycystic kidney disease</td>
<td>7</td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Glomerulopathies</td>
<td>6</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Haemolytic and uremic syndrome</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Metabolic disease</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Post-chemotherapy</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>8</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Total=238</td>
<td>106</td>
<td>105</td>
<td>23</td>
<td>4</td>
</tr>
</tbody>
</table>

Legend: Patients’ renal disorders and CKD classification. CKD: chronic kidney disease, CKD stage I, II, III, IV and V denote GFR ≥ 90, 60-89, 30-59, 15-29 and < 15 ml/mn per 1.73 m², respectively.