Relationship of gender, age, and body mass index to errors in predicted kidney function

Massimo Cirillo, Pietro Anastasio and Natale G. De Santo

Nephrology, Second University of Naples (SUN), Naples, Italy

Abstract

Background. Previous studies have shown conflicting data on accuracy of equations for kidney function prediction. The present work analysed the relationship of gender, age and body mass index (BMI) to error of predictions by the Cockcroft–Gault equation (CGeq), the simplified equation of the Modification of Renal Diseases Study (MDRD eq) and the Mayo Clinic equation (Mayoeq).

Methods. Inulin clearance (glomerular filtration rate; GFR) and other variables were measured in 380 subjects of both sexes, aged 18–88 years, with and without kidney disease. GFR was defined as low when < 60 ml/min × 1.73 m². BMI was used for definition of underweight/overweight. Relative error of predictions was used as an index of bias. It was calculated as prediction minus GFR (positive values = overestimates, negative values = underestimates) and expressed as a percentage of the GFR. Absolute error was used as an index of imprecision and was calculated as the absolute value of relative error.

Results. CGeq relative error was inversely associated with age and directly associated with BMI (P < 0.001), but not with gender or GFR. MDRD eq relative error was inversely associated with female gender and GFR (P < 0.001), but not with age or BMI. Mayoeq relative error was directly associated with male gender, BMI and GFR (P < 0.01), but not with age. Absolute error was higher for CGeq than for MDRD eq but only at low GFR (P < 0.001). Mayoeq had a higher absolute error than CGeq and MDRD eq (P < 0.01).

Conclusions. Errors of predictions varied not only with GFR but also with gender, age and BMI. Without using creatinine assay calibration, Mayoeq was less accurate than both MDRD eq and CGeq, whereas MDRD eq was slightly more precise than CGeq but only at low GFR.

Keywords: Cockcroft–Gault equation; creatinine; glomerular filtration rate; MDRD equation

Introduction

Glomerular filtration rate (GFR) is considered to be the key marker of kidney function [1,2]. However, the precise measurement of GFR is invasive, time-consuming, expensive and technically difficult [2]. There is a strong, non-linear, inverse relationship between GFR and the serum concentration of creatinine, which is a skeletal muscle catabolite [2]. Thus, for a given GFR, differences in serum creatinine reflect differences in muscular mass [2]. Equations have been developed to predict kidney function by using serum creatinine data combined with readily available indices of muscular mass, such as gender, age and weight [2].

The most common equations used in adults are the Cockcroft–Gault equation (CGeq) and the simplified equation from the Modification of Diet in Renal Disease Study (MDRD eq) [3–5]. These two equations differ in many aspects, including the predicted index (creatinine clearance for CGeq, true GFR for MDRD eq), units of predictions (ml/min for CGeq, ml/min × 1.73 m² for MDRD eq) and variables used for the prediction (serum creatinine, gender, age and body weight for CGeq; serum creatinine, gender, age and race for MDRD eq). According to the data of Levey et al. [4], the MDRD eq is more accurate than the CGeq. However, data from other laboratories did not unanimously support this conclusion [6–12]. Coresh et al. [13] reported that differences in the creatinine assay lower the accuracy of the MDRD eq, and suggested that creatinine assays should be calibrated with that of the MDRD laboratory for use in the MDRD eq [13]. None of the previous studies that examined accuracy of prediction equations was performed with use of this calibration [14]. Rule et al. [11] reported low accuracy of the MDRD eq in healthy persons examined at the...
Mayo Clinic [11], and proposed a new equation for use in healthy individuals and in kidney disease (Mayoeq) [12].

Previous reports on prediction equations focused mainly on differences between persons with and without kidney dysfunction [6–12]. However, no studies have examined whether accuracy of equations may differ between men and women, or between young and older ages, even though gender and age are key determinants of the predictions. There are also no data on the possible role of underweight and overweight, even though a difference between the equations is the inclusion of body weight data in the CGeq but not in the MDRDeq or Mayoeq.

In the present study, we investigated CGeq, MDRDeq and Mayoeq while using inulin clearance as the gold standard in individuals with and without kidney dysfunction. We aimed to analyse the effects of gender, age and underweight/overweight on the error of predictions.

Methods

Individuals, measurements and definitions

Individuals aged ≥18 years were enrolled prospectively in the study beginning from 1997 and after the first report of Levey et al. [15] that examined prediction equations. During selection, the patients underwent GFR measurement by inulin clearance and routine medical investigation with inclusion of urinalysis and serum creatinine measurement. Individuals with and without kidney disease were included in the study. However, individuals having treatment with drugs affecting creatinine secretion were excluded.

Inulin clearance was measured in the morning after an overnight fast as previously reported [16]. Briefly, individuals were given a bladder catheter for urine collection and a venous cannula in each arm (for inulin infusion and collection of blood samples, respectively). Inulin infusion was initiated as a bolus and continued at a constant rate to maintain plasma inulin at ~20 mg/dl. After initiation of the constant rate infusion, the test included a 90 min period for equilibration and two consecutive 45 min measurement periods. A light hydration was maintained throughout these periods by administration per os of tap water (0.5 ml/kg every 30 min). Separate urine collections were obtained for each period. During equilibration, plasma inulin was measured at 30, 60 and 90 min. Individuals were excluded from the analysis if plasma inulin variations exceeded ±4 mg/dl at 60 or 90 min (unstable plasma inulin). During the 45 min measurement periods, blood samples were withdrawn at initiation and at completion of urine collections. Inulin clearance was calculated separately for the two 45 min periods as urinary excretion rate/average plasma concentration (between initiation and completion of urine collection). The average of inulin clearances between the two 45 min periods was used for statistical analyses. The intra-individual variation of inulin clearance between the two periods was 8.07% of the mean value. Venous blood samples were collected before initiation of inulin infusion and used for measurements of serum creatinine by a kinetic alkaline picrate assay that was tested daily for internal and external quality control (high values >1.40 mg/dl). The intra-individual variation of serum creatinine between the two blind measurements was 4.68% of the mean value.

The prediction of creatinine clearance by the CGeq (ml/min) was calculated as 140–age × weight[kg]/72 × serum creatinine[mg/dl] with use of the 0.850 multiplier for female gender as per the reported method [3]. Predictions by the CGeq were normalized per 1.73 m² of body surface area (BSA) for comparison with predictions of the other equations. BSA (m²) was calculated as 0.007184 × heightm/0.72 × weightkg [17]. The expected bias of the CGeq predictions was an overestimate of measured GFR due to tubular creatinine secretion [2,18]. Thus, the main analyses for CGeq predictions were done with use of the correction factor reported by Levey et al. [4] (corrected CGeq predictions = CGeq predictions/1.16). The prediction of GFR by the MDRDeq (ml/min×1.73 m²) was calculated as 186 × serum creatinine[1.134][mg/dl] × age[0.203], with the use of the 0.742 multiplier for female gender as per the reported method [5]. The term for black individuals in the MDRDeq was not used since all individuals were white. The equation was used without creatinine assay calibration from the MDRD laboratory. Thus, the MDRDeq predictions could be affected by a bias due to differences in the creatinine assay between the MDRD laboratory and our laboratory [13]. The prediction of GFR by the Mayoeq (ml/min×1.73 m²) was calculated as exp [1.911 + 5.249/serum creatinine[mg/dl] – 2.114/serum creatinine[10.205] – 0.00686 × age – 0.205 (if female)], with serum creatinine values <0.8 mg/dl set to 0.8 mg/dl as per the reported method [12].

Individuals were defined as having low GFR when inulin clearance was <60 ml/min×1.73 m² and as having non-low GFR when inulin clearance was ≥60 ml/min×1.73 m². Persons were defined as free of kidney disease when they reported no kidney disease, when GFR was not low and in the absence of proteinuria and haematuria at urinalysis. Body mass index (BMI = weight[kg]/heightm²) was used for analyses of underweight and overweight. Four BMI strata were used for the analyses: ≤21 kg/m² (underweight), 21–24.99 kg/m² (normal weight), 25–29.99 kg/m² (overweight) and ≥30 kg/m² (obesity). Diabetes mellitus was defined as fasting serum glucose ≥126 mg/dl (≥7 mmol/l) and/or treatment with insulin and/or antidiabetic drugs.

Statistics

Inulin clearance was used as the gold standard measurement of GFR. Relative error and absolute error were used as indices of accuracy. The relative error was calculated as the prediction minus the measured GFR. It indicated the tendency of the prediction to overestimate or to underestimate the measured GFR (positive relative error for prediction > measured GFR, negative relative error for prediction < measured GFR). The absolute error was calculated as the absolute value of the prediction minus measured GFR. It indicated the amplitude of the error (imprecision), regardless of the tendency to overestimate or underestimate. Errors were expressed as a percentage of measured GFR to reduce the effects of different use of GFRs upon the error (large errors at high GFR, small errors at low GFR). The expression of errors as a percentage of measured GFR removed the effect of BSA normalization to affect the accuracy of the CGeq predictions (see Appendix).
Correlates of error in predicted kidney function

**Results**

**Descriptive statistics**

The study cohort was comprised of 380 individuals (age range 18–88 years): 178 with previous diagnosis of renal disease (n = 38 for glomerular disease, n = 24 for nephroangiosclerosis, n = 20 for diabetic nephropathy, n = 15 for polycystic kidney disease, n = 8 for pyelonephritis, n = 73 for other or unknown diseases), 150 with previous diagnosis of non-renal disease (undergoing inulin clearance measurements as part of diagnostic or therapeutic protocols) and 52 not reporting any disease (healthy volunteers or candidate kidney donors).

The main analyses for indices of kidney function were with data normalized per 1.73 m² of BSA, since measured GFR without BSA normalization correlated with BSA in persons with low and non-low GFR (R = 0.462 and 0.438, P < 0.001). Figure 1 shows correlations of individual data of measured GFR with predictions by CGeq, MDRDeq and Mayoq. Table 1 reports descriptive data for gender, age, indices of kidney function and other variables.

Age was inversely correlated with measured GFR and with GFR from predictions from all equations, but only in analyses limited to persons without kidney disease (R > 0.319, P < 0.001). Figure 2 shows means of measured GFR and means of measured GFR and of predictions stratified by age in persons without kidney disease. The regression line of measured GFR over age had a slope of −0.44 ml/min × 1.73 m² per year (95% CI = −0.63−0.26). The slopes were not significantly different in the MDRDeq and Mayoq (−0.48 and −0.68 ml/min × 1.73 m² per year), but were significantly more negative for corrected CGeq ($R = −0.97 / C2$ × 1.73 m² per year, 95% CI = $−1.17−0.76$). Findings were similar when persons with BMI ≥30 kg/m² were excluded (not shown).

**Accuracy for entire cohort and by GFR level**

Table 2 reports means of relative and absolute error for the entire cohort and for subgroups with low and

**Table 1.** Descriptive statistics: prevalence or median (min/max)

<table>
<thead>
<tr>
<th>Gender, n</th>
<th>214/166</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), n</td>
<td></td>
</tr>
<tr>
<td>18-24</td>
<td>33</td>
</tr>
<tr>
<td>25-34</td>
<td>48</td>
</tr>
<tr>
<td>35-44</td>
<td>93</td>
</tr>
<tr>
<td>45-54</td>
<td>99</td>
</tr>
<tr>
<td>55-64</td>
<td>72</td>
</tr>
<tr>
<td>65-88</td>
<td>35</td>
</tr>
</tbody>
</table>

Kidney function, ml/min × 1.73 m² of BSA

<table>
<thead>
<tr>
<th>Prediction</th>
<th>75.9 (8.2/158.8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGeq prediction</td>
<td>79.1* (7.9/199.0)</td>
</tr>
<tr>
<td>Corrected CGeq prediction</td>
<td>68.2** (6.8/171.5)</td>
</tr>
<tr>
<td>MDRDeq prediction</td>
<td>73.9** (5.0/165.0)</td>
</tr>
<tr>
<td>Mayoq prediction</td>
<td>97.7* (6.8/155.4)</td>
</tr>
<tr>
<td>With low GFR&lt;sup&gt;a&lt;/sup&gt;, n</td>
<td>149</td>
</tr>
<tr>
<td>Without kidney disease&lt;sup&gt;b&lt;/sup&gt;, n</td>
<td>198</td>
</tr>
<tr>
<td>Serum creatinine, mg/dl&lt;sup&gt;d&lt;/sup&gt;</td>
<td>1.00 (0.51/9.02)</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>70.0 (44.5/167.0)</td>
</tr>
<tr>
<td>Body surface area, m²</td>
<td>1.78 (1.34/2.62)</td>
</tr>
</tbody>
</table>

Individuals by BMI stratum, n

| BMI |  
| --- | --- |
| BMI <21 kg/m² | 45 |
| BMI 21–24.99 kg/m² | 128 |
| BMI 25–29.99 kg/m² | 116 |
| BMI ≥30 kg/m² | 91 |
| With diabetes mellitus, n | 37 |

<sup>a</sup>Corrected for overestimate of measured GFR by creatinine clearance (see text).

<sup>b</sup>With GFR <60 ml/min × 1.73 m².

<sup>c</sup>Without reported kidney disease, low GFR, proteinuria and haematuria.

<sup>d</sup>To convert mg/dl to μmol/l, multiply by 88.4.

<sup>e</sup>Higher vs measured GFR (P < 0.01, paired t-test with logarithm-transformed data to control for skewed distribution).

<sup>f</sup>Lower vs measured GFR (P < 0.01, paired t-test with logarithm-transformed data to control for skewed distribution).

---

#### Figures

**Fig. 1.** Individual data and identity line of measured GFR (inulin clearance) correlated with predictions by CGeq (left panel), MDRDeq (middle panel) and Mayoq (right panel). Open symbols are for individuals without kidney disease (without reported kidney disease, low GFR, proteinuria and haematuria). Correlation coefficients between predictions and measured GFR were 0.806 for CGeq, 0.866 for MDRDeq and 0.882 for Mayoq.

**Table 2** reports means of relative and absolute error for the entire cohort and for subgroups with low and...
Absolute error of corrected CG\textsubscript{eq} predictions was similar for the entire cohort and for the subgroups with low and non-low GFR (Table 2). MDRD\textsubscript{eq} predictions had lower absolute error compared with corrected CG\textsubscript{eq} predictions in analyses for the entire cohort due only to a difference in the subgroup with low GFR. Mayo\textsubscript{eq} predictions had the highest absolute error in all analyses. The difference was significant vs CG\textsubscript{eq} predictions and MDRD\textsubscript{eq} predictions in analyses for the entire cohort and for the low GFR subgroup, but not for the non-low GFR subgroup. Prevalence of absolute error <30% was lower for CG\textsubscript{eq} than for MDRD\textsubscript{eq} (67.6 vs 71.6%, \(P = 0.097\) by McNemar test). Mayo\textsubscript{eq} predictions had lower prevalence of absolute error <30% compared with MDRD\textsubscript{eq} predictions (59.5 vs 71.6%, \(P < 0.001\)) and with CG\textsubscript{eq} predictions (59.5 vs 67.6%, \(P = 0.010\)).

Gender and error of predictions

Figure 3 shows relative error data for the entire study cohort stratified by gender. The relative error of corrected CG\textsubscript{eq} predictions was not significantly different between men and women (similar findings for non-corrected CG\textsubscript{eq} predictions; data shown in the legend of Figure 3). In contrast, the relative errors of predictions by the MDRD\textsubscript{eq} and Mayo\textsubscript{eq} were significantly different between men and women (an underestimate in women for MDRD\textsubscript{eq} and an overestimate in men for Mayo\textsubscript{eq}). This gender effect was consistent at low and non-low GFRs for all equations. For corrected CG\textsubscript{eq} predictions, the relative error was similar in men and women with low and non-low GFR (data not shown). For MDRD\textsubscript{eq} predictions, the relative error was more negative in women than men in the low GFR subgroup (–12.7 vs +3.9%, \(P < 0.001\)) and in the non-low GFR subgroup (–14.4 vs –7.6%, \(P = 0.137\)). For Mayo\textsubscript{eq} predictions, the relative error was more negative (or less positive) in women than men in the low GFR subgroup (–5.1 vs –1.9%, \(P = 0.561\)) and in the non-low GFR subgroup (+9.1 vs 25.9%, \(P < 0.001\)).

The absolute error of predictions was similar between men and women for CG\textsubscript{eq} (23.4 vs 25.1%, \(P = 0.370\)), higher in women than men for MDRD\textsubscript{eq} (22.5 vs 20.6%, \(P = 0.093\)) and higher in men than women for Mayo\textsubscript{eq} (31.0 vs 23.1%, \(P < 0.001\)).

Age and error of predictions

Figure 4 shows relative error data for the entire study cohort stratified by age. The relative error of corrected CG\textsubscript{eq} predictions varied linearly over age from overestimates in young adults to underestimates in older adults (similar findings for non-corrected CG\textsubscript{eq} predictions; data shown in the legend of Figure 4). The relative error of predictions by MDRD\textsubscript{eq} and Mayo\textsubscript{eq} did not vary linearly with age. Findings on the effect of age were consistent at low and non-low GFR for all equations. The relative error of corrected CG\textsubscript{eq} predictions varied linearly with age in the low GFR subgroup (95% CI = –8.2/–2.2%), for the low GFR subgroup (95% CI = –8.1/–0.5%) and for the non-low GFR subgroup (95% CI = –11.4/–2.0%). The relative error of MDRD\textsubscript{eq} predictions was an underestimate for the entire cohort (95% CI = –8.7/–3.5%), which was smaller at low GFR (95% CI = –6.0/–0.2%) than at non-low GFR (95% CI = –15.2/–6.2%). The relative error of Mayo\textsubscript{eq} predictions was an overestimate for the entire cohort (95% CI = +7.0/+13.6%) due to combination of an overestimate for the non-low GFR subgroup (95% CI = +15.4/+22.2%) and an error not significantly different from zero for the low GFR subgroup (95% CI = –9.4/+2.2%).

![Fig. 2. Means of measured GFR and of predictions for corrected CG\textsubscript{eq}, MDRD\textsubscript{eq} and Mayo\textsubscript{eq} stratified by age in individuals without kidney disease (without reported kidney disease, low GFR, proteinuria and haematuria; \(n = 198\)). \(P \leq 0.001\) from ANOVA among age strata for all indices (\(n \text{ by age stratum} = 22, 24, 62, 59, 22 \text{ and } 9\)).](https://academic.oup.com/ndt/article-abstract/20/9/1791/1850343/1791)
The absolute error of predictions was higher in persons at the extreme of age distribution than in those in the middle for CGeq (ages 18–24 and 65–88 combined together vs ages 25–64 = 28.0 vs 23.3%, \( P = 0.045 \)), but not for MDRD\textsubscript{eq} (21.8 vs 21.3%, \( P = 0.793 \)) or Mayo\textsubscript{eq} (28.9 vs 27.3%, \( P < 0.520 \)).

**BMI and error of predictions**

Figure 5 shows relative error data for the entire study cohort stratified by BMI. The relative error of corrected CG\textsubscript{eq} predictions varied linearly over BMI from underestimates in underweight to overestimates in overweight (similar findings for non-corrected CG\textsubscript{eq} predictions; data shown in the legend of Figure 5). The relative error of MDRD\textsubscript{eq} predictions was negative for all BMI values and did not vary significantly from underweight to obesity. The relative error of Mayo\textsubscript{eq} predictions was positive for all BMI strata and varied from slightly positive values in underweight to largely positive values in obesity. These BMI findings were consistent at low and non-low GFR for CG\textsubscript{eq} and MDRD\textsubscript{eq}, but not for Mayo\textsubscript{eq}. For corrected CG\textsubscript{eq} predictions, the relative error varied linearly with BMI in the low GFR subgroup (from \(-15.7\% \) at BMI <21 to \(+3.8\% \) at BMI \( \geq 30, \ P = 0.018 \)) and in the non-low GFR subgroup (from \(-22.2\% \) at BMI <21 to \(+23.5\% \) at BMI \( \geq 30, \ P < 0.001 \)). For MDRD\textsubscript{eq} predictions, the relative error was not associated with BMI in the low GFR subgroup nor in the non-low GFR subgroup (data not shown, \( P > 0.3 \)).

For Mayo\textsubscript{eq} predictions, the relative error varied linearly with BMI in the non-low GFR subgroup (from \(+10.1\% \) at BMI <21 to \(+28.2\% \) at BMI \( \geq 30, \ P = 0.002 \)), but not in the low GFR subgroup (from \(-9.3\% \) at BMI <21 to \(-2.4\% \) at BMI \( \geq 30, \ P = 0.561 \)).
The absolute error of predictions was higher in obese than in non-obese subjects for CGeq (28.8 vs 22.7%, \( P=0.004 \)) and Mayoeq (32.8 vs 25.6%, \( P=0.004 \)), but not for MDRDeq (20.0 vs 21.9%, \( P=0.303 \)).

The differences in relative error among the BMI subgroups could be due to the effect of weight which is a determinant of BMI. This possibility was investigated by dividing the study cohort into tertiles of weight by using a BMI-stratified procedure to create subgroups having different weights (means 69.9, 73.6 and 80.3 kg, \( P<0.001 \)) but similar BMI (27.9, 27.9 and 27.4 kg/m\(^2\), \( P=0.247 \)). In the groups with different weights and similar BMI, the test for a linear trend in the relative error was not significant for corrected CGeq predictions (\(-3.8, -11.1\) and \(-0.8\), \( P=0.147 \)), MDRDeq predictions (\(-2.9, -10.5\) and \(-5.0\), \( P=0.494 \)) or Mayoeq predictions (+12.6, +3.3 and +13.9\%, \( P=0.784 \)).

**Discussion**

The present study was the first to examine the effects of gender, age and BMI on errors in kidney function predictions as assessed by CGeq, MDRDeq and Mayoeq. The relative error (bias) of CGeq predictions was associated with age and BMI, but not with gender and GFR. CGeq predictions tended to overestimate measured GFR in obesity, and underestimate GFR in underweight and older ages. MDRDeq predictions showed lower average values than measured GFR. This bias was explained by an underestimate of measured GFR in female gender and in non-low GFR. Mayoeq predictions showed higher average values than measured GFR mainly due to overestimates in male gender and in obesity. These findings tended to be similar at low and non-low GFR. Thus, effects of gender, age and BMI on error of predictions were independent of GFR level, with the exception of the association between BMI and error of Mayoeq predictions which was found only in subjects with non-low GFR. In the present study, GFR level affected the error of predictions by MDRDeq and Mayoeq, but not by CGeq. Although predictions by both MDRDeq and Mayoeq had greater error at non-low GFR than at low GFR, these errors were opposite, with underestimates using the MDRDeq, and overestimates using the Mayoeq. Absolute error was high in subgroups with large relative error (underestimate or overestimate). Thus, predictions had high absolute error (low precision) in obesity and in young/older ages while using CGeq in female gender and non-low GFR while using MDRDeq, and in male gender, non-low GFR and obesity while using Mayoeq. Study limitations included the absence of data for various ethnic groups, the size of the study cohort, and the absence of creatinine calibration. The absence of creatinine calibration, which is needed for the MDRDeq, the CGeq and the Mayoeq, may limit comparisons between our results and other laboratories since differences in creatinine assays affect the error of predictions [13].

**Gender and error of predictions**

Equations for prediction of kidney function include a correction factor that lowers the prediction in women to compensate for the gender-dependent difference in creatinine generation (i.e. muscular mass). The relative error of predictions in the present study was similar in men and women using the CGeq, was 13% higher in men than in women using the MDRDeq, and was 10% higher in men than in women using the Mayoeq. Thus, the difference in muscular mass between men and women in the present cohort was adequately predicted by the CGeq but was overestimated by the MDRDeq and the Mayoeq. It is unlikely that data for the gender effect on the error of predictions were entirely explainable by confounding of creatinine assay and/or GFR level since findings were similar at low and non-low GFR. According to an alternative possibility, gender may be a determinant of inaccurate predictions.
since a coefficient derived in a given cohort may not adequately predict gender-associated differences in muscular mass in another cohort. In fact, gender-dependent differences in muscular mass may vary among cohorts of different studies depending on ethnicity, anthropometrics, age distribution, prevalence of healthy and diseased individuals, and a diet or drug treatment. This interpretation is in accordance with the heterogeneity of factors proposed for female gender in the equations examined in the present study (0.85 for the CGeq, 0.72 for the MDRDeq and 0.81 for the Mayoeq).

**Age and error of predictions**

Equations for prediction of kidney function include a coefficient for age that corrects for the ageing-associated decline in creatinine generation (muscular mass). In the present study, the relative error was not significantly different by age strata with MDRDeq and Mayoeq, but was different by age strata with the CGeq prediction. Because these findings were similar at low and non-low GFR, they were independent of a major confounding due to creatinine assay and/or GFR level. To explain these findings, it is possible that the age coefficient of the CGeq overestimated the ageing-associated decline in muscular mass. This finding may have been specific for the cohort examined in the present study. This possibility seems unlikely since the lack of association between age and error of predictions by MDRDeq and Mayoeq suggests similar ageing-associated declines in muscular mass among the subjects examined in the present study, in the MDRD study and in the study of the Mayo Clinic.

**BMI and error of predictions**

In contrast to the MDRDeq and Mayoeq, the CGeq includes a coefficient for body weight to correct the prediction of inter-individual differences in creatinine generation (muscular mass) due to body mass. We found that relative errors of predictions were significantly associated with BMI for predictions with CGeq and Mayoeq, but not with MDRDeq.

For CGeq, the relative error was linearly associated with BMI, with a large overestimate in obesity. Because these findings were similar at low and non-low GFR, it was unlikely that they were explained only by confounding of creatinine assay and/or GFR level. They may be explained by the rule that weight coefficients of the CGeq do not differentiate between muscular mass (relevant to creatinine generation) and non-muscular mass (not relevant to creatinine generation). Thus, the CGeq transforms any weight difference into a difference in predicted kidney function and tends to overestimate kidney function in overweight. The possibility that the data merely reflected a cohort-specific phenomenon is unlikely since a previous report showed a population-based association between BMI and CGeq predictions. In fact, BMI of middle-aged participants in the ARIC study was significantly higher (+5.1 kg/m²) in persons with non-low CGeq predictions than in persons with low CGeq predictions [19]. Previously, Cockcroft and Gault suggested that a correction to lean weight should be used for the CGeq in the presence of excess fat; however, they did not report data showing the effects of obesity on predictions [3]. A correction to lean weight for the CGeq is rarely used and would require additional validation and calculations to adjust for gender, age, health/disease and other factors. For the Mayoeq, the relative error was progressively more positive along BMI strata only in persons with non-low GFR. It is unlikely that this finding is explained by the creatinine assay. Instead, the overweight/obese subjects with non-low GFR in the present study had lower muscular mass than the overweight/obese subjects with non-low GFR in the Mayo study.

**Conclusions**

The CGeq and the MDRDeq are commonly tools for prediction of kidney function. Previous studies have demonstrated that the error of predictions is sensitive to creatinine assay calibration and GFR level. Although the effect of creatinine assay can be reduced by cross-calibration of the assay among different laboratories, this procedure has not been used [14], and this is not possible for old equations such as the CGeq. The results of the present study showed that variability in predictions of kidney function depend on gender, age and BMI. Thus, conflicting data in previous studies regarding accuracy of equations may have reflected differences in the distribution of gender, age and BMI among the cohorts in the studies. In the present study, performed without calibration of creatinine assay, precision was slightly higher for MDRDeq than for CGeq at low GFR, and similar for CGeq and MDRDeq at non-low GFR. The equation recently developed in the Mayo Clinic (Mayoeq) was the least precise in the present study. Therefore, a quadratic equation by itself does not improve the accuracy of predictions based on non-calibrated creatinine assay. The small differences in precision between the CGeq and MDRDeq found in the present study are of limited practical importance because of the large effect of other confounders, such as errors in serum creatinine measurements, variability in muscular mass (creatinine generation) and tubular creatinine secretion. More importantly, our results suggest that the accuracy of predictions tends to decrease in older age and/or in obesity for the CGeq, and in non-low GFR for MDRDeq.

**Conflict of interest statement.** None declared.

**References**


**Appendix**

It has been hypothesized that the BSA normalization of data affected analyses of error of CGeq predictions since the CGeq was designed for predictions expressed as ml/min [20]. The expression of errors as a percentage of measured GFR excluded this possibility since the percentage error of a prediction calculated without BSA normalization is the same as the percentage error calculated with BSA normalization. For example, the error is $+10.0\%$ in a person with a CGeq prediction of 110.0 ml/min and a measured GFR of 100.0 ml/min. If the BSA of a subject was 1.602 m$^2$, the BSA normalization would change the CGeq prediction up to 118.8 ml/min $/1.73m^2$ and the measured GFR up to 108.0 ml/min $/1.73m^2$. The error is $+10.0\%$ also for these BSA normalized data. The stability of the error indicates that during BSA normalization, CGeq prediction and measured GFR are corrected for the same multiplier, which is 1.73 m$^2$/individual BSA (for the case in the example: 1.730/1.602 = 1.080).