

$$
\sqrt{\text{ILEY}}
$$

Towards equations for estimating glomerular filtration rate without demographic characteristics

Dear Editor,

Estimated glomerular filtration rate (eGFR) is an essential index for identifying and monitoring chronic kidney disease (CKD). For estimating GFR, the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) research group recently proposed two new serum creatinine-based equations without race, eGFRcr(AS) and eGFRcr-cys(AS), to replace the widely used equations eGFRcr(ASR) and eGFRcr-cys(ASR), that considers serum creatinine, cystatin C (CysC), age, sex and race. $1-4$ The new equations were developed and validated using data mainly from White and Black participants.⁴ It remains unknown whether they are applicable to plasma samples as well as to non-White and non-Black subjects. The present study addressed these two questions with a focus on two Chinese CKD cohorts and showed that equations based on other blood metabolites could perform better.

The two cohorts comprised 52 patients from Guangzhou (Third Affiliated Hospital) and 135 patients from Macau (Kiang Wu Hospital) for whom serum samples and plasma samples were available, respectively (Tables S1 and S2 for their characteristics). Performance of the equations was examined with reference to the GFR measured by iohexol plasma clearance (mGFR). $2,4$ The bias and precision of the current and new CKD-EPI equations were similar for the two cohorts (Table [1\)](#page-1-0). The accuracy of all the equations was better for the Macau cohort, as revealed by lower values of the root mean square error and higher values of the percent agreement within 30% of the mGFR (P_{30}) , the percent agreement within 20% of the mGFR (P_{20}) and the percent agreement between the mGFR and eGFR categories.⁵ As plasma samples were used in the Macau cohort, our results suggested that both old and new CKD-EPI equations should be applicable to plasma samples. In the subsequent analyses, the data from the two Chinese cohorts were pooled for a larger sample size.

A *P*³⁰ value of 80%–90% is considered acceptable, and a P_{30} value of 90% or higher is preferred.⁴ However, P_{30}

values of two new equations without race were less than 80% for the two Chinese cohorts (Table 1, Result S1 for statistical test results). LOWESS curve fitting of the agreement between the mGFR and eGFR indicated that all the serum creatinine-based equations overestimated the GFR in the range of 30–90 ml/min/1.73 m² (Figure [1\)](#page-3-0). Overestimation was the highest for the eGFRcr(AS) equation (Figure [1B\)](#page-3-0), whereas it was not obviously for the eGFR- $\cos(AS)$ equation (Figure [1E\)](#page-3-0). In a recent study, applying the eGFRcr(AS) equation to the White European population also resulted in shifting a major proportion of CKD patients to a higher eGFR category.⁶ In addition to the unsatisfactory P_{30} values, the use of the new equations could lead to underestimation of the disease severity in Chinese patients with mild or moderate loss of kidney function. Therefore, we recommend not to use the new equations without race for Chinese CKD patients.

To minimize the effect of systemic and systematic biases, $6,7$ to ensure a metabolite marker can be applied to both serum and plasma specimens, 8 and to avoid falsepositive biomarkers, 9 we adopted a stringent two-centre study design (Figure [2\)](#page-4-0), involving plasma samples from 10 healthy volunteers in addition to the two Chinese cohorts (Table S3, Discussion S1 about study designs). The samples were subjected to untargeted metabolomics profiling using a Metabolon's Discovery HD4 platform. 10 The final list of putative biomarkers contained 212 metabolites (207 negatively and 5 positively correlated with the mGFR, Figure [2](#page-4-0) and Result S2 for details). Among the top 20 putative biomarkers (Table S5), hydroxyasparagine and *N,N*-dimethyl-proline–proline had not been previously shown to be inversely associated with kidney function (Table S6, Discussion S2 about accessibility in clinical practice). Furthermore, one of the top 20 putative biomarkers was creatinine, indicating the success of our study design.

For proof-of-concept, we attempted to construct eGFR equations from the data of the top 20 putative biomarkers in the discovery dataset. *S*-Adenosylhomocysteine,

This is an open access article under the terms of the [Creative Commons Attribution](http://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

^{© 2022} The Authors. *Clinical and Translational Medicine* published by John Wiley & Sons Australia, Ltd on behalf of Shanghai Institute of Clinical Bioinformatics.

TABLE 1 Summary of the performance of different Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equations for estimating the GFR in Chinese chronic kidney disease (CKD) patients using the Guangzhou cohort (serum samples) and Macau cohort (plasma samples)

(Continues)

TABLE 1 (Continued)

Abbreviations: eGFR, estimated glomerular filtration rate; mGFR, measured glomerular filtration rate.

^aThe performance of each equation in estimating the GFR was evaluated in terms of bias, precision and accuracy, according to the methods described by Inker et al. $(2012)^2$ $(2012)^2$ $(2012)^2$ and Inker et al. $(2021)^4$.

^bThe values indicating the best two equations were italicized and made bold.

 c Accuracy was calculated as the RMSE relative to the mGFR, the percentage of estimates that differed from the mGFR by less than 30% (P_{30}), and the percentage that differed by less than 20% (P_{20}).

gluconate, N6-succinyladenosine and hydroxyasparagine (*p* values <.05) were retained in the final regression equation named CKD-msMET4a (Table [2\)](#page-5-0). As gluconate is a common component of food additives and supplements, the performance of CKD-msMET4a could be affected by gluconate intake. We hence constructed another equation named CKD-msMET4b with the exclusion of gluconate. The major difference from the first equation was the replacement of gluconate with gulonate (Result S3 for additional information). Gulonate and gluconate are stereoisomers of each other, indicating that the two equations are almost identical in nature. Further inclusion of

age, sex, height, weight and/or serum/plasma CysC did not significantly improve the two equations (*p* values >.05).

Using the data from the independent validation cases, the performance of the two equations was compared with that of the CKD-EPI eGFR equations (Figure [2\)](#page-4-0). Concerning bias, precision, accuracy and GFR category agreement, either CKD-msMET4a or CKD-msMET4b equations appeared to be one of the best two equations although the statistical power was insufficient (Table [2,](#page-5-0) Result S4 for detailed results). The P_{30} values of both equations were 82%. They were not significantly different from those of the current eGFRcr-cys(ASR) equation (*p*-values >.890)

FIGURE 1 Comparison of the measured glomerular filtration rate (mGFR) and estimated glomerular filtration rate (eGFR) for alternative GFR estimating equations in the two Chinese cohorts. The Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equations for estimating the GFR are named according to the name provided by Inker et al.^{[4](#page-7-0)} They are referred to by the filtration marker or markers (creatinine [eGFRcr], cystatin C [eGFRcys] or creatinine–cystatin C [eGFRcr-cys]) and the demographic factors (age, sex and race [ASR] or age and sex [AS]) that were used in their development. They are (A) eGFRcr(ASR), (B) eGFRcr(AS), (C) eGFRcr-cys(ASR), (D) eGFRcr-cys(AS), and (E) eGFRcys(AS). Data from the two Chinese chronic kidney disease (CKD) cohorts (Guangzhou cohort and Macau cohort) are shown. Overall, we referred to 185 CKD patients from the 2 cohorts. Bias was defined as the mean difference (median difference) between the mGFR and eGFR. A positive sign indicates an overestimation of the mGFR, and a negative sign indicates an underestimation of the mGFR. P_{30} is the percent agreement within 30% of the mGFR. GFR category agreement is the percent agreement between the mGFR and eGFR categories (<15, 15–29, 30–59, 60–89 and \geq 90 ml/min/1.73 m²). For each plot, the LOWESS curve was fitted using the data from 185 CKD patients.

but were significantly different from the P_{30} value of the new eGFRcr(AS) equation (*p*-values <.05). LOWESS curves revealed that the overestimation of GFR in the range of 30–90 ml/min/1.73 $m²$ was not observed for the CKD-msMET4a and CKD-msMET4b equations (Figure S1, Result S4 for detailed results).

Although the male and female patients might have differences in age and total mass muscles, GFRs estimated by the two equations were not significantly different between the male and female patients (*p*-values <.005, Table S7, Result S4 for detailed results). This suggests that four

metabolites could provide sufficient information for estimating the GFR without demographic characteristics. This may also suggest that these two equations are applicable to patients of other races, such as the White and Black populations.

In conclusion, the new CKD-EPI creatinine-based equations without race could lead to underestimation of the disease severity in Chinese patients with mild or moderate loss of kidney function. However, our results suggest the possibility of developing new metabolite equations for estimating the GFR without demographic characteristics.

CLINICAL AND TRANSLATIONAL MEDICINE

FIGURE 2 Schematic illustration of the two-centre study design and participant flow for discovery and validation of the metabolite markers for estimating the GFR in patients with chronic kidney disease (CKD). The study design comprised two parts. Part 1 was for biomarker discovery and construction of equations for estimating the GFR, whereas Part 2 was for independent validation of the putative biomarkers and equations. The study design involved (i) two types of blood samples, serum and plasma, from two centres and (ii) plasma samples from healthy volunteers to minimize false discovery of biomarkers associated with systemic and/or systematic bias(es). The serum samples were collected from 52 Chinese CKD patients at the Third Affiliated Hospital of Sun Yat-sen University in Guangzhou; the plasma samples were collected from 135 Chinese CKD patients at Kiang Wu Hospital and 10 healthy volunteers in Macau, resulting in a total of 197 subjects. Seventy-nine CKD patients (19 CKD patients from Guangzhou and 60 CKD patients from Macau, accounting for 40% of all the participants) were randomly selected to form an independent validation group. The rest of the participants were assigned for biomarker discovery and for construction of equations to estimate the GFR. The statistical methods used at different steps and the summary of the corresponding results are also indicated in the figure. Before the partial correlation analysis and equation construction, the measured glomerular filtration rate (mGFR), normalized abundances of the putative metabolite markers, serum/plasma creatinine level and serum/plasma cystatin-C (cysC) level were logarithmically transformed. Top 20 candidates of 215 putative biomarkers (pseudouridine, *N*-acetylserine, C-glycosyltryptophan, 3-(3-amino-3-carboxypropyl)uridine, 5,6-dihydrouridine, hydroxyasparagine, N6-carbamoylthreonyladenosine, *S*-adenosylhomocysteine, gulonate, 2,3-dihydroxy-5-methylthio-4-pentenoate, creatinine, *N*-acetylneuraminate, *N,N*-dimethyl-proline–proline, erythronate, N6-succinyladenosine, *N*-acetylalanine, *O*-sulfo-l-tyrosine, arabonate/xylonate, 4-acetamidobutanoate and gluconate) were used to construct equations for estimating GFR. Equation 1, CKD-msMET4a is as follows: estimated glomerular filtration rate ($eGFR$) = $exp(3.999 - .137)$ ln(normalized abundance of *S*-adenosylhomocysteine) − .162 ln(normalized abundance of gluconate) − .209 ln(normalized abundance of N6-succinyladenosine) − .319 ln(normalized abundance of hydroxyasparagine)). Equation 2, CKD-msMET4b is as follows: eGFR = exp(3.991 − .141 ln(normalized abundance of *S*-adenosylhomocysteine) − .146 ln(normalized abundance of gulonate) − .214 ln(normalized abundance of N6-succinyladenosine) − .270 ln(normalized abundance of hydroxyasparagine)). *Gluconate and gulonate are stereoisomers of each other, indicating that the two equations, CKD-msMET4a and CKD-msMET4b, are almost identical in nature. The two equations have performance similar to that of the current Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) creatinine-CysC equation with race, that is eGFRcr-cys(ASR).

eGFRcr-cys(AS) [new]

TABLE 2 Summary of the performance of different equations for estimating the GFR in the independent validation group^a

I

(Continues)

TABLE 2 (Continued)

Performance evaluation^b – equations for estimating the GFR, ml/min/1.73 m² of body-surface area

Abbreviations: CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; mGFR, measured glomerular filtration rate.

aIndependent validation cases included serum samples from 19 CKD patients, the Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou and plasma samples from 60 CKD patients, Kiang Wu Hospital, Macau.

^bThe performance of each equation in estimating the GFR was evaluated in terms of bias, precision and accuracy, according to the methods described by Inker et al. $(2012)^2$ $(2012)^2$ $(2012)^2$ and Inker et al. $(2021)^4$.

cAlthough the statistical power was insufficient, the values suggesting the best two equations were italicized and made bold

dCKD-msMET4a equation, eGFR = exp(3.999 − .137 ln(normalized abundance of *S*-adenosylhomocysteine) − .162ln(normalized abundance of gluconate) − .209 ln(normalized abundance of N6-succinyladenosine) − .319 ln(normalized abundance of hydroxyasparagine)).

eCKD-msMET4b equation, eGFR = exp(3.991 − .141 ln(normalized abundance of *S*-adenosylhomocysteine) − .146 ln(normalized abundance of gulonate) − .214 ln(normalized abundance of N6-succinyladenosine) − .270 ln(normalized abundance of hydroxyasparagine)).

 $\rm ^f$ Accuracy was calculated as the RMSE relative to the mGFR, the percentage of estimates that differed from the measured GFR by less than 30% (P_{30}), and the percentage that differed by less than 20% (P_{20}).

ACKNOWLEDGEMENTS

This work was supported by The Macao Science and Technology Development Fund (FDCT) (Grant Number: 0032/2018/A1). We gratefully acknowledge the contribution of all members of field staff who were involved in planning and conducting this study. We thank all participants in our studies for their donation of blood and time. We also thank the Calibra-Metabolon Joint Laboratory (Hangzhou, China) for their support. Finally, we thank Yuanyan Tang for support with the study data. This work was also supported by the Faculty of Health Sciences of the University of Macau and Kiang Wu Hospital Charitable Association, Macau.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The datasets used and analysed during the current study are available from the corresponding authors on reasonable request.

Hongquan Peng¹ Irene Ling Ang² Xun Liu $³$ </sup> Chiwa Aoieong1 Tou Tou¹ Tsungyang Tsai¹ Kamleong Ngai4 Hao I. Cheang⁴ Peijia Liu³ Terence Chuen Wai Poon²

1 *Department of Nephrology, Kiang Wu Hospital, Macau, China*

²*Pilot Laboratory, Institute of Translational Medicine, Centre for Precision Medicine Research and Training, Faculty of Health Sciences, University of Macau, Macau, China*

³*Department of Nephrology, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, China* ⁴*Clinical Laboratory, Kiang Wu Hospital, Macau, China*

Correspondence

Hongquan PENG, Department of Nephrology, Kiang Wu Hospital, Macau, China. Email: hpeng93170@gmail.com

Terence Chuen Wai Poon, Pilot Laboratory, Institute of Translational Medicine, Centre for Precision Medicine Research and Training, Faculty of Health Sciences, University of Macau, Macau, China. Email: tcwpoon@um.edu.mo

ORCID

Hongquan Peng <https://orcid.org/0000-0003-4893-7903> *Terence ChuenWai Poon* [https://orcid.org/0000-0003-](https://orcid.org/0000-0003-4723-6910) [4723-6910](https://orcid.org/0000-0003-4723-6910)

REFERENCES

- 1. Levey AS, Coresh J, Tighiouart H, Greene T, Inker LA. Measured and estimated glomerular filtration rate: current status and future directions. *Nat Rev Nephrol*. 2019;16(1):51-64.
- 2. Inker LA, Schmid CH, Tighiouart H, et al. Estimating glomerular filtration rate from serum creatinine and cystatin C. *N Engl J Med*. 2012;367(1):20-29.
- 3. Gama RM, Kalyesubula R, Fabian J, Mahalingasivam V. NICE takes ethnicity out of estimating kidney function. *BMJ*. 2021;374:n2159.
- 4. Inker LA, Eneanya ND, Coresh J, et al. New creatinine- and cystatin C-based equations to estimate GFR without race. *N Engl J Med*. 2021;385(19):1737-1749.
- 5. Kidney disease:Improving global outcomes (KDIGO) CKD work group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int Suppl*. 2013;3:1-150.
- 6. Fu EL, Coresh J, Grams ME, et al. Removing race from the CKD-EPI equation and its impact on prognosis in a predominantly White European population. *Nephrol Dial Transplant*. 2022:gfac197.
- 7. Chen C, Gowda GAN, Zhu J, et al. Altered metabolite levels and correlations in patients with colorectal cancer and polyps detected using seemingly unrelated regression analysis. *Metabolomics*. 2017;13(11):125.
- 8. Kaluarachchi M, Boulangé CL, Karaman I, et al. A comparison of human serum and plasma metabolites using untargeted (1)H NMR spectroscopy and UPLC-MS. *Metabolomics*. 2018;14(3): 32.
- 9. Zhang P, Ang IL, Lam MMT, et al. Susceptibility to false discovery in biomarker research using liquid chromatography-high resolution mass spectrometry based untargeted metabolomics profiling. *Clin Transl Med*. 2021;11(6):e469.
- 10. Long T, Hicks M, Yu HC, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nat Genet*. 2017;49(4):568-578.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.