Is There Something Better than the Best Marker of Kidney Function?

Brad C. Astor* and Nrupen A. Bhavsar†‡

*Departments of Medicine and Population Health Sciences, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin; †Section of Value and Comparative Effectiveness, Division of General Internal Medicine, Department of Medicine, New York University School of Medicine, New York, New York; and ‡Division of General Internal Medicine, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland


Measurement of kidney function is critical both for clinical decision-making and research. Direct measurement of GFR using urinary or plasma clearance of exogenously administered markers, such as 125I-iothalamate, is considered the gold standard method to assess kidney function. These procedures, however, are burdensome and impractical in many settings. Using endogenous markers to estimate GFR is an attractive alternative, and serum creatinine-based estimates are widely reported by clinical laboratories.

The potential biases and inaccuracies of serum creatinine-based GFR estimates are widely known and include variations in creatinine generation as a result of differences in muscle mass or diet and variations in tubular secretion or extrarenal elimination of creatinine. Although serum levels of cystatin C are less affected by muscle mass, other nonrenal determinants of serum cystatin C have been identified. The limitations of directly measured GFR are less well studied but are significant and include substantial measurement error, dietary protein intake, and day-to-day and diurnal variations. The inaccuracies and biases of the different methods to assess GFR have specific implications relevant to each of its uses.

Hsu et al. compared the cross-sectional associations between GFR measured with 125I-iothalamate (iGFR) and estimated GFR based on serum creatinine (eGFR-Cr) or cystatin C (eGFR-cysC) with chronic kidney disease (CKD)-associated complications in a subset of 1214 participants in the Chronic Renal Insufficiency Cohort (CRIC) Study. Inclusion criteria were an eGFR-Cr between 20 and 70 ml/min per 1.73 m² for those aged 21 to 44 years, 20 and 60 ml/min per 1.73 m² for those aged 45 to 64 years, and 20 and 50 ml/min per 1.73 m² for those aged 65 to 74 years. Individuals with polycystic kidney disease, kidney transplant recipients, and those with severe comorbid conditions, such as advanced heart failure, were excluded. The authors assessed the linear associations of iGFR, eGFR-Cr, and eGFR-cysC with levels of hemoglobin, potassium, bicarbonate, and phosphate, as well as with dichotomized outcomes based on these values.

As expected, lower iGFR or eGFR was associated with lower levels of hemoglobin and bicarbonate and higher levels of potassium and phosphorus. The correlations, however, were relatively weak, with R² values ranging from 0.07 to 0.13. Similarly, logistic regression analyses of dichotomized outcomes resulted in C statistics ranging from 0.67 to 0.73, indicating only moderate ability to discriminate between participants with and without each complication. The authors point out that iGFR resulted in slightly better discrimination than eGFR-Cr or eGFR-cysC for anemia but slightly poorer discrimination for hyperphosphatemia, although these differences were not tested statistically. Discrimination for metabolic acidosis and hyperkalemia were similarly weak across all three GFR measures.

The CRIC Study used rigorous data collection methods, including a detailed protocol for iGFR measurements and high-quality laboratory methods. The subset of study participants included in these analyses, consisting of approximately one third of all CRIC Study participants, is representative of the overall CRIC Study population. The results presented provide important information with implications relevant to both clinical treatment and research.

There are several elements of the study that warrant discussion. All three GFR measures were poorly correlated with all four CKD complications. Although few studies have reported R² or C statistics from linear or logistic regression models, as was done here, the correlations observed by Hsu et al. are somewhat lower than those found in similar populations in several previous studies, although the reasons for these differences are unknown. Although one would not expect a linear relationship between GFR and complications in stages 2 through 4 CKD, the authors found that nonlinear models, including splines, did not substantially improve the fit of the models in this study. Such nonlinearities would be difficult to detect statistically in the presence of such weak associations.

It is important to note that the coefficient of variation (CV) of iGFR was much higher (13.8%) than that for creatinine (1.1%) or cystatin C (4.9%) for any of the outcomes studied (0.9 to 4.5%). This relatively poor reliability was found despite the CRIC Study’s efforts to minimize variation...
in GFR as a result of extrarenal factors by regulating diet, posture, intake of nonsteroidal anti-inflammatory agents, and time of day of the GFR measurements. Similar reliability for two iGFR measurements taken a mean of 62 days apart was reported for the African American Study of Kidney Disease and Hypertension (AASK) and the Modification of Diet in Renal Disease (MDRD) Study. The variability associated with iGFR measurements, even under these rigorous study protocols, limits the precision with which a single iGFR measurement can be used to assess steady-state kidney function.

The current findings are in agreement with other studies of cross-sectional associations and with longitudinal studies investigating prediction of mortality, cardiovascular outcomes, and kidney failure, which have found equal or stronger associations with eGFR compared with iGFR. They call into question the status of iGFR as the gold standard measurements of kidney function. There are, however, potential alternative explanations for the findings that must be considered, most notably confounding of the observed associations by nonrenal determinants.

For example, inflammation is associated with both decreased creatinine production (and therefore higher eGFR-Cr) and with lower hemoglobin. This would have the effect of reducing any observed association between lower eGFR-Cr and anemia. Similarly, higher dietary intake of red meat is associated with higher iGFR and higher serum creatinine but also increased creatinine production (and therefore higher eGFR-Cr) and phosphorus away from the null. This confounding would move any observed association between lower iGFR and phosphorus toward the null and move any observed association between lower eGFR-Cr and phosphorus away from the null.

The authors pose the question, “Is iGFR really better than eGFR?” To address this question, one must ask, better for what purpose? For diagnosis, the best estimate of true GFR may be most useful. Likewise, the best estimate of true change in GFR is most useful for clinically monitoring of treatments intended to preserve kidney function or that are potentially nephrotoxic. It remains unclear which methods are best in either scenario. Other factors, such as body habitus or comorbid conditions, may dictate the answer, which is likely to be some combination of multiple markers.

For prognosis, however, the most accurate GFR measurement may not be most informative measurement. In the Third National Health and Nutrition Examination Survey (NHANES III), GFR estimated from both creatinine and cystatin C, which more closely approximates iGFR than GFR estimated from either marker alone, was significantly less predictive of outcomes than eGFR-cysC. These results demonstrate that the best estimate of iGFR is not necessarily the best predictor of outcomes.

The results of this study are most applicable to guiding selection of individuals for screening for CKD-related complications. In this case, as with prognosis, the best measure of filtration may not be the most accurate predictor of which individuals are most likely to benefit from screening and subsequent treatment. Specific situations may require a different balance between the relatively poor precision with iGFR against the potential biases of endogenous markers. Again, a multiple marker approach may minimize the influence of specific confounders of any single marker while providing a more reliable estimate of true GFR.

We agree with the authors’ statement that we may not be able to identify a single measure that provides a sufficient overall index of kidney function. Finding the optimal approaches for incorporating measures of kidney function into decision-making for diagnosis, screening, monitoring, and testing new interventions, however, remains a worthy goal that will require even more work.

DISCLOSURES
None.

REFERENCES


See related article, “Measured GFR Does Not Outperform Estimated GFR in Predicting CKD-Related Complications,” on pages ●●●●●●●●.