## Item S2. Supplementary Methods

GFRs estimated on the same subject using different equations would likely be correlated because most of the equations use similar variables. Since all the metrics used for assessing equation performance are based on estimated GFR, therefore these metrics would also be correlated.

We used generalized estimating equations (GEE) to circumvent the problem of non-independence of the variables. This approach assumes that measurements on different subjects are independent while the measurements on the same subjects are correlated. In the model, we used the subject as the grouping variable, and used an unstructured correlational matrix to account for within subject correlation of the variables. This approach provided robust standard errors.

For the patient level variables of $A_{30}$ (binary variables representing if the absolute percent error is less than or equal to 30%) the model obtained the odds of success of the eGFR equations.

$$\text{Odds of Success} = {P_{30}}/{(1 - P_{30})} \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{Equation 1}$$

$$P_{30} = {\text{Odds of Success}}/{(1+\text{Odds of successs})} \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{Equation 2}$$

We estimated the $P_{30}$ from the odds of success (Equations 1 & 2). We also calculated the difference between the probability of success ($P_{30}$) of the CKD-EPI or the MDRD Study equation and the alternative equations as well as their 99.6 % CIs by running 2000 simulation by using methods described previously.[1]

The null hypothesis was that the difference between the $P_{30}$ of the CKD-EPI or the MDRD Study equation is < 0 with an alternative hypothesis that the difference is >0 (Equation 3).

$$\text{H0: } P_{30}^{\text{CKD-EPI or MDRD}} - P_{30}^{\text{Alternative Equation}} \leq 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{Equation 3}$$

$$\text{HA: } P_{\mathbf{30}}^{\text{CKD-EPI or MDRD}} - P_{30}^{\text{Alternative Equation}} > 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots Equation 3$$

We used GEE models to obtain the estimated difference between the mean absolute bias of the CKE-EPI or the MDRD Study equations and the alternative equations. The null hypothesis was that the mean difference was greater than or equal to 0 with an alternative hypothesis that it was less than 0 (Equation 4).

$$\text{H0: Mean Absolute Error}^{\text{CKD-EPI or the MDRD}} - \text{Mean Absolute Error}^{\text{Alternative Eqauation}} \geq 0.. \text{Equation 4}$$

$$\text{H0: Mean Absolute Error}^{\text{CKD-EPI or the MDRD}} - \text{Mean Absolute Error}^{\text{Alternative Eqauation}} < 0.. Equation 4$$

Descriptive Statistics: GFR and error were normally distributed; therefore we estimated the mean GFR and error by a linear GEE model (gaussian distribution). Percent error and absolute percent error were not normally distributed on the relative scale therefore we reported median percent error and median absolute error for these metrics. Since the CKD-EPI and the MDRD study equations were derived to minimize the error between log(mGFR) and log(eGFR), for comparisons between these equations, we also calculated the mean percent error and mean percent absolute error on the log scale [1 - exp[$\Sigma$ { log(eGFR)-log(mGFR) } / n ]].

Patient level covariates: We estimated $P_{30}$ of each equation in previously described subgroups by adding interaction terms as categorical or continuous variables in the GEE models to assess their effect on equation performance. We obtained the global p values for the interactions terms from the model. A significant p value (<0.01) indicates that the difference between the performance of the CKD-EPI or the MDRD Study equation and the alternative equation varies significantly across the levels of the categories based on patient characteristics. We explored this further by comparing the performance of the CKD-EPI or the MDRD Study equations with the alternative equations in each category graphically (forest plots of the difference of the $P_{30}$ of the CKD-EPI or the MDRD Study equation and each of the alternative equations for each characteristic). If the upper limit of the 99.6% confidence interval (CI) failed to cross 0, the alternative equation is superior to the CKD-EPI or the MDRD Study equation; if the CI included 0, the performance of the equations is similar while if the lower limit of the CI crossed 0, the CKD-EPI or the MDRD Study equation is superior to the alternative equation in that category (Figures S2-S7).

**Item S2 Reference**

1. King G, Tomz M, Wittenberg J. Making the Most of Statistical Analyses: Improving Interpretation and Presentation. *American Journal of Political Science*. 2000;44(2):341–355.