INTRODUCTION

Recently, the misuse of race as a biological variable, instead of a social construct, in biomedical research and in clinical decision-making tools used to estimate kidney function and inform clinical decisions related to timing of transplant referral, evaluation, wait listing, and approval of living kidney donors and allocation of kidneys has received national attention. Specifically, contemporary kidney function estimation tools founded on erroneous, historic beliefs that the dichotomization of Black versus non-Black in a correction factor contained meaningful biologic information has led to an underestimation of chronic kidney disease burden among Black patients and contributed to delayed referral and wait listing of Black patients with kidney transplant patients (December 2008–February 2020) that collects patient-reported race. We computed Cohen’s kappa statistic to estimate agreement between provider-perceived and patient-reported race in the 2 data sources. We used an unadjusted generalized linear model to examine changes in agreement over time. Results. Among 2942 kidney transplant patients, there was almost perfect agreement among Asian (kappa = 0.88, 95% confidence interval [CI], 0.84–0.92), Black (kappa = 0.97, 95% CI, 0.96–0.98), and White categories (kappa = 0.95, 95% CI, 0.93–0.96) and worse agreement among Hispanic/Latino (kappa = 0.66, 95% CI, 0.57–0.74) and Native Hawaiian/Other Pacific Islander categories (kappa = 0.40, 95% CI, 0.01–0.78). The percent agreement decreased over time (difference in percent agreement = −0.55, 95% CI, −0.75 to −0.34). However, there were differences in these trends by race: −0.07/y, 95% CI, −0.21 to 0.07 for Asian; −0.06/y, 95% CI, −0.28 to 0.16 for Black; −0.01/y, 95% CI, −0.21 to 0.19 for Hispanic/Latino; −0.43/y, 95% CI, −0.58 to −0.28 for White categories. Conclusions. Race misclassification has likely led to increasingly biased research estimates over time, especially for Asian, Hispanic/Latino, and Native Hawaiian/Other Pacific Islander study populations. Improvements to race measurement include mandating patient-reported race, expanding race categories to better reflect contemporary US demographics, and allowing write-ins on data collection forms, as well as supplementing data with qualitative interviews or validated measures of cultural identity, ancestry, and discrimination.
chronic kidney disease. Much of the discourse around race correction in kidney function estimation has focused on social construct definition and the consequences of imprecision. Less commonly addressed, however, is measurement imprecision of race.

Race is a multidimensional, relational social construct that is often collected by self-report or perceived by an observer. More specifically, provider-perceived race is a social assignment that is based on a variable combination of physical appearance (eg, skin color, hair texture, nose, and lip shape) and interpersonal interactions (eg, language, surname, attire) and varies according to observer demographics and familiarity with different racial groups. Perceived race is associated with differential exposure to racism, with Hispanic and American Indian individuals who are perceived to be White and lighter-skinned Black individuals reporting higher health status than those who are perceived to be their actual self-identified race or are darker-skinned. Self-reported race is how one identifies when presented with closed-ended survey instruments and is recommended in the Office of Management and Budget’s Standards for the Classification of Federal Data on Race and Ethnicity and used in vital statistics and population-based research. Although race is routinely present in kidney transplant (KT) registry data and commonly used in research, race measurement according to its unique dimensions has not been examined. It is possible that agreement between provider-perceived race and self-reported race may differ across commonly used data sources used to study KT patients.

Therefore, the goal of this study was to examine race measurement in kidney transplantation research. We used national, transplant registry data and data from a prospective, multicenter cohort of adult KT candidates and recipients (1) to quantify race agreement between provider-perceived race and patient self-reported race and (2) to examine the change in this agreement over time according to race classification.

PATIENTS AND METHODS

Study Population

This study leveraged 2 prospective cohorts of 2443 adult (18 y and older) patients with end-stage renal disease who were waitlisted for a KT and 1244 adult living and deceased donor KT recipients at Johns Hopkins Hospital, Baltimore, MD (candidates enrolled November 2009–February 2020, n=2443; recipients enrolled December 2008–February 2020, n=1244) and the University of Michigan University Hospital (recipients enrolled March 2015–November 2016, n=83). After excluding duplicate patient records (ie, patients who were waitlisted and then transplanted), a total of 2942 patients with complete records were included in the analysis.

We linked each patient record to the Scientific Registry of Transplant Recipients (SRTR). The SRTR data system includes data on all donors, waitlisted candidates, and transplant recipients in the United States, submitted by the members of the Organ Procurement and Transplantation Network (OPTN). The Health Resources and Services Administration, US Department of Health and Human Services, provides oversight to the activities of the OPTN and SRTR contractors. This dataset has previously been described elsewhere.

All research activities being reported are consistent with the Declaration of Helsinki and the Declaration of Istanbul. The study protocol was approved by the Johns Hopkins University Institutional Review Board and the University of Michigan Institutional Review Board. All cohort study participants provided written informed consent.

Classification of Race and Ethnicity

At cohort enrollment, baseline characteristics and year were collected using a standard data collection form. The patient’s self-reported race was ascertained using a multiple-choice question with options that included White, Black, Asian/Other, Native Hawaiian/Other Pacific Islander, and Other. An additional question was used to assess Hispanic ethnicity. For both race and ethnicity questions, patients were permitted to select 1 option only.

Transplant registry data on race is derived from OPTN data collection forms that are used by transplant hospitals, histocompatibility laboratories, and Organ Procurement Organizations. The OPTN collects provider-reported race and ethnicity in 1 question and permits the selection of >1 option. SRTR converts this data into race categories that include White, Black/African American, American Indian/Alaska Native, Asian, Native Hawaiian/Other Pacific Islander, Arab/Middle Eastern, Indian Subcontinent, Hispanic/Latino, and Other and a separate ethnicity category that includes categories Hispanic/Latino and Non-Hispanic-Latino. In this study, we considered patients categorized as White, Black, Black/African American, American Indian/Alaska Native, Asian, Native Hawaiian/Other Pacific Islander to be non-Hispanic. We reclassified patients who were American Indian/Alaska Native, Arab/Middle Eastern, or Indian Subcontinent in the SRTR database into other to make congruent with cohort data for analysis.

Analysis

For analysis, the construct of race collected in the prospective cohorts was considered “patient self-reported race and ethnicity,” and that reported in the transplant registry was considered “provider-perceived race and ethnicity.” We computed the Cohen kappa statistic to estimate the agreement beyond expected by chance. A kappa coefficient of 0 indicates chance agreement, whereas 1 indicates perfect agreement, specifically ≤0 poor, 0.01 to 0.20 slight, 0.21 to 0.40 fair, 0.41 to 0.60 moderate, 0.61 to 0.80 substantial, and 0.81 to 0.99 almost perfect agreement. We then estimated the percent agreement between patient’s self-reported race and provider-perceived race for the overall study period using a pairwise Pearson correlation. After visually checking the distribution of observed agreement over time, we used unadjusted generalized linear models to assess the mean change in agreement per year.

All analyses were performed using Stata, version 15 (StataCorp, College Station, TX). Two-sided P values <0.05 were considered statistically significant.

RESULTS

We identified 2942 patients, of which 81.1% (n=2385) were waitlisted candidates and 18.9% (n=557) were KT recipients. Of the total population, 41.1% (n=120) were Asian, 42.3% (n=1245) were Black, 3.4% (n=101) were Hispanic/Latino, 0.2% (6) were Native Hawaiian or Other Pacific Islander, and 47.7% (n=1404) were White individuals.
The agreement decreased 0.55% per year (difference in percent linear regression model, we found that the percent observed data collection year (r = –0.86, with a strong correlation between percent agreement and provider-perceived race declined linearly over time (Figure 1), CI, –0.01 to 0.17) (Table 1).

For the overall study population, the kappa agreement was almost perfect (kappa = 0.92, 95% confidence interval [CI], 0.91-0.93). After stratifying by race category, we found that the kappa agreement was almost perfect for Asian (kappa = 0.88, 95% CI, 0.84-0.92), Black (kappa = 0.97, 95% CI, 0.96-0.98), and White (kappa = 0.95, 95% CI, 0.93-0.96) populations. The kappa agreement was substantial for the Hispanic population (kappa = 0.66, 95% CI, 0.57-0.74), fair for Native Hawaiian or Other Pacific Islander (kappa = 0.40, 95% CI, 0.01-0.78), and slight for Other (kappa = 0.08, 95% CI, –0.01 to 0.17) (Table 1).

Agreement between patient self-reported race and provider-perceived race declined linearly over time (Figure 1), with a strong correlation between percent agreement and data collection year (r = –0.86, P < 0.001). Using an unadjusted linear regression model, we found that the percent observed agreement decreased 0.55% per year (difference in percent agreement = –0.55, 95% CI, –0.75 to –0.34). According to race, the changes in percent agreement per year were –0.07% per year for Asian (difference in percent agreement = –0.07, 95% CI, –0.21 to 0.07), –0.06% per year for Black (difference in percent agreement = –0.06, 95% CI, –0.28 to 0.16), –0.01% per year for Hispanic/Latino (difference in percent agreement = –0.01, 95% CI, –0.21 to 0.19), and –0.43% per year for White (difference in percent agreement = –0.43, 95% CI, –0.58 to –0.28) categories (Figure 2).

**DISCUSSION**

In this study comparing prospective, multicenter cohort data to national transplant registry data, we found racial differences in agreement between patient self-reported and provider-perceived race, with almost perfect agreement being observed among Asian, Black, and White populations and worse agreement among Hispanic/Latino and Native Hawaiian/Other Pacific Islander populations. We also found that agreement between patient self-reported and provider-perceived race decreased linearly over time.

Our finding of relatively higher race misclassification among patients who are not classified as White when race from administrative data sources compared with self-reported race is consistent with prior studies. In a study comparing Medicare claims race data to self-reported race collected during a home healthcare visit, race was misclassified for 7.5% of the overall population and was greatest among Hispanic populations.

Racial misclassification has also been found to vary across settings, with better agreement being observed in geographic regions with high concentrations of racial and ethnic minority individuals. In states with high concentrations of Hispanic populations, race variable imputation using surname improved race classification from 7.5% to 69.4% as compared with administrative data alone. Prior studies incorporating large, commonly used administrative data sources, such as Medicare claims and Surveillance, Epidemiology, and End Results program, have found misclassification to be associated with immigration status, preferred language, marital status with presumed surname changes, and educational attainment.

We also found a novel finding that race misclassification has been worsening over time. Temporal trends in race misclassification could reflect the increasing number of multiracial US citizens and decreasing proportion who identify as White alone, changes in US political or social context over

**TABLE 1.** Agreement between patient self-reported race and ethnicity and provider-perceived race and ethnicity in kidney trans-plant patients (n = 2942) between 2009 and 2020

| Race/ethnicity               | Patient self-reported race prevalence (%) | Provider-perceived race prevalence (%) | Observed agreement | Kappa coefficient (95% CI) |
|------------------------------|-------------------------------------------|---------------------------------------|--------------------|---------------------------|
| Overall                      | –                                         | –                                     | 95.3               | 0.92 (0.91-0.93)          |
| Asian                        | 4.1                                       | 4.8                                   | 99.0               | 0.88 (0.84-0.92)          |
| Black                        | 42.3                                      | 43.2                                  | 98.6               | 0.97 (0.96-0.98)          |
| Hispanic                     | 3.4                                       | 2.3                                   | 98.1               | 0.66 (0.57-0.74)          |
| Native Hawaiian or other Pacific Islander | 0.2                                      | 0.1                                   | 99.8               | 0.40 (0.01-0.78)          |
| White                        | 47.7                                      | 49.5                                  | 97.3               | 0.95 (0.93-0.96)          |
| Other                        | 2.2                                       | 0.1                                   | 97.8               | 0.08 (–0.01 to 0.17)      |

Self-reported race/ethnicity was collected in a prospective cohort study of kidney transplant candidates and recipients. Provider-perceived race/ethnicity was obtained from the Scientific Registry of Transplant Recipients. Cohen’s kappa coefficients with 95% CIs were presented.

CI, confidence interval.

![Lowess plot](image-url)

**FIGURE 1.** Observed agreement between patient self-reported race and ethnicity and provider-perceived race and ethnicity in kidney transplant patients (n = 2942) between 2009 and 2020. The relationship between agreement and year was examined by Lowess plot.
time affecting patients’ willingness to identify as part of certain racial groups, or outdated data collection procedures that inappropriately combine patients from unique social hierarchies. One of the most recently debated inaccuracies is the classification of Middle Eastern and North African (MENA) as White in Census data and other administrative data sources. In a study using an online, crowdsourcing platform to obtain a nonprobability of sample respondents and conduct a factorial experiment, Maghbouleh found that respondents who identify as MENA change their responses based on survey options, with 88% identifying as MENA alone when given the option.24

Our work is not without limitations. Although it is certain that race is self-reported in the prospective cohort, it is not certain whether transplant centers require patients to confirm their preferred race classification for inclusion in the national transplant registry, and this is an insurmountable limitation in this retrospective study. It is likely that race ascertainment in the registry is a combination of perceived and self-reported race. Given that this nondifferential measurement error biases toward the null, our findings are likely an overestimation and racial misclassification is more problematic than reported. Both data sources used in this study collected race using closed-ended survey instruments, which can differ from open-ended, self-identified race. We treated self-reported race as the gold standard for data collection, but there is no gold standard race measurement. Both self-reported race and perceived race are valid social constructs for applicable research hypotheses.8 For example, perceived race is important for research questions concerning the influence of discrimination on health outcomes, whereas self-identified race is more important for understanding social networks, behaviors, and attitudes.8 We included multcenter, prospective cohort data that included an “Other” category in data collection, which is uninformative and inconsistent with contemporary guidance on the reporting of race and ethnicity in medical and science journals. The “Other” category in the cohort study likely captures American Indian or Alaska Native and/or multiracial populations, and prior research has demonstrated racial misclassification with geographic variation like what is observed among Hispanic/Latino populations. 20,21 The absence of this population in our study, though unfortunate and contributing to historic marginalization, does not influence our findings and is a target for improvement and investigation in future research. Finally, it is possible that our finding of less-than-perfect concordance between the 2 data sources may also reflect data entry errors rather than different race measurement procedures. Notwithstanding, we are likely underestimating the magnitude of information bias present in kidney transplantation research.

Our work is important because race misclassification in national data in kidney transplantation has likely introduced information bias into surveillance reports and research evaluating access to and outcomes of transplantation. Prior studies have found that race misclassification in national, administrative data sources underestimates cancer cases by 16% to 37% for Asians18 and cancer incidence in Native Americans20 and

**FIGURE 2.** Observed agreements between patient self-reported race and ethnicity and provider-perceived race and ethnicity over time during 2009 to 2020. For each race/ethnicity, the observed agreements by year of candidate data collection in the national registry were depicted as the scatter points, and the trend over time was presented as a fitted line.
distorts the magnitude of racial disparities in maternal and child health indicators used in national health statistic reports. Prior studies have also found qualitative differences in effect estimates after adjustment for misclassification. Although misclassification of race is not unexpected given its dynamic and inherently imperfect social construction that hinges upon person, place, and time, there are strategies that can be used to mitigate information bias. Moving forward, improvements in using registry-based race variable might include acknowledging the limitations of race definition and categorization in manuscripts, expanding race and ethnicity categories beyond federal administrative standards to reflect contemporary US demographics, including sensitivity analyses accounting for race misclassification, supplementing data with interviews to allow for patient self-report, including a write-in line on data collection forms, or including validated measures more specific to the research hypothesis of how race influences health outcomes such as skin tone, cultural identity, geographic ancestry, and everyday racism or discrimination experiences.

In this study comparing race measurement in national, transplant registry data to a multicenter, prospective cohort collecting patient self-reported race, we found that race classification was different across the 2 data sources for the same patients. Misclassification of race has likely led to increasingly biased kidney transplantation research estimates over time, especially for studies including populations who are classified as Asian, Hispanic/Latino, or Native Hawaiian/Other Pacific Islander.

ACKNOWLEDGMENTS

The data reported here have been supplied by the Hennepin Healthcare Research Institute as the contractor for the SRTR. The interpretation and reporting of these data are the responsibility of the author(s) and in no way should be seen as an official policy of or interpretation by the SRTR or the US Government.

REFERENCES

1. Vyasa DA, Eisenstein LG, Jones DS. Hidden in plain sight—reconsidering the use of race correction in clinical algorithms. N Engl J Med. 2020;383:874–882.
2. Chu CD, Powe NR, Crews DC, et al. CKD progression from the time of estimated GFR-based waitlist eligibility and racial disparities in transplant access. Am J Kidney Dis. 2022;79:841–848.e1.
3. Hoening MP, Mann A, Pavlakis M. Removal of the Black race coefficient from the estimated glomerular filtration equation improves transplant eligibility for Black patients at a single center. Clin Transplant. 2022;36:e14467.
4. Zeinick LR, Leca N, Young B, et al. Association of the estimated glomerular filtration rate with vs without a coefficient for race with time to eligibility for kidney transplant. JAMA Nephrol. 2021;4:e2034004.
5. Ahmed S, Nutt CT, Eneanya ND, et al. Examining the potential impact of race multiplier utilization in estimated glomerular filtration rate calculation on African-American care outcomes. J Gen Intern Med. 2021;36:464–471.
6. Tsai JW, Cerdeña JP, Goedel WC, et al. Evaluating the impact and rationale of race-specific estimations of kidney function: estimations from U.S. NHANES, 2015–2016. Eclinicalmedicine. 2021;42:101197.
7. Doshi MD, Singh N, Hippen BE, et al. Transplant clinician opinions on use of race in the estimation of glomerular filtration rate. Clin J Am Soc Nephrol. 2021;16:1552–1559.
8. Roth WD. The multiple dimensions of race. Ethn Racial Stud. 2016;39:1310–1338.
9. Harris DR. In the eye of the beholder: observed race and observer characteristics. PSC Research Report No. 02-522. August 2002. Available at https://psc.irr.umich.edu/pubs/pdf/rr02-522.pdf. Accessed June 24, 2022.
10. Jones CR, Truman BI, Elam-Evans LD, et al. Using “socially assigned race” to probe White advantages in health status. Ethn Dis. 2008;18:496–504.
11. Cobb RJ, Thomas CS, Laster Pirtle WN, et al. Self-identified race, socially assigned skin tone, and adult physiological dysregulation: Assessing multiple dimensions of “race” in health disparities research. SSM Popul Health. 2016;2:595–602.
12. Klionoff EA, Landrine H. Is skin color a marker for racial discrimination? Explaining the skin color-hypertension relationship. J Behav Med. 2000;23:329–338.
13. Monk EP Jr. The cost of color: skin color, discrimination, and health among African-Americans. AJS. 2015;121:396–444.
14. Office of Management and Budget. Revisions to the standards for the classification of federal data on race and ethnicity. October 30, 1997. The White House: President Barack Obama. Available at https://obamawhitehouse.archives.gov/omb/fedreg_1997standards. Accessed June 24, 2022.
15. Massie AB, Kucikra LM, Segev DL. Big data in organ transplantation: registries and administrative claims. Am J Transplant. 2014;14:1729–1790.
16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–174.
17. Grafowa IB, Jarrin OF. Beyond black and white: mapping misclassification of medicare beneficiaries race and ethnicity. Med Care Res Rev. 2021;78:616–626.
18. Gormez SL, Glaser SL. Misclassification of race/ethnicity in a population-based cancer registry (United States). Cancer Causes Control. 2006;17:771–781.
19. Jarrin OF, Nyandenge AN, Grafowa IB, et al. Validity of race and ethnicity codes in medicare administrative data compared with gold-standard self-reported race collected during routine home health care visits. Med Care. 2020;58:e1–e5.
20. Frost F, Taylor V, Fries E. Racial misclassification of Native Americans in a surveillance, epidemiology, and end results cancer registry. J Natl Cancer Inst. 1992;84:957–962.
21. Jim MA, Arias E, Senerca DI, et al. Racial misclassification of American Indians and Alaska Natives by Indian Health Service Contract Health Service Delivery Area. Am J Public Health. 2014;104 Suppl 3:S295–S302.
22. Swallen KC, West DW, Steward SL, et al. Predictors of misclassification of Hispanic ethnicity in a population-based cancer registry. Ann Epidemiol. 1997;7:200–206.
23. Jones N, Marks R, Ramirez R, et al. 2020 census illuminates racial and ethnic composition of the country. August 12, 2021. Available at https://www.census.gov/library/stories/2021/08/improved-race-ethnicity-measures-reveal-united-states-population-much-more-multiracial.html. Accessed March 15 2022.
24. Maghbouleh N, Schachter A, Flores RD. Middle Eastern and North African Americans may not be perceived, nor perceive themselves, to be White. Proc Natl Acad Sci U S A. 2022;119:e2117940119.
25. Buescher PA, Gizlice Z, Jones-Vessey KA. Discrepancies between published data on racial classification and self-reported race: evidence from the 2002 North Carolina live birth records. Public Health Rep. 2005;120:393–398.
26. Klonof EA, Landrine H. Is skin color a marker for racial discrimination? Explaining the skin color-hypertension relationship. Biometrics. 1998;54:774–781.
27. Garcia JA, Sanchez GR, Sanchez-Youngman S, et al. RACE AS LIVED EXPERIENCE: The impact of multi-dimensional measures of race/ethnicity on the self-reported health status of Latinos. Du Bois Rev. 2018;32:349–373.
28. Bonham VL, Green ED, Pérez-Stable EJ. Examining how race, ethnicity, and ancestry data are used in biomedical research. JAMA. 2018;320:1533–1534.