



HHS Public Access

Author manuscript

Am J Prev Med. Author manuscript; available in PMC 2024 July 01.

Published in final edited form as:

Am J Prev Med. 2024 January ; 66(1): 182–184. doi:10.1016/j.amepre.2023.08.019.

Reporting of Race and Ethnicity in the Vaccine Safety Datalink, 2011–2022

Kate E. Kurlandsky, BA¹, Amy B. Stein, PhD¹, Simon J. Hambidge, MD, PhD¹, Eric S. Weintraub, MPH², Joshua T.B. Williams, MD¹

¹Ambulatory Care Services, Denver Health, Denver, Colorado;

²Immunization Safety Office, Centers for Disease Control and Prevention, Atlanta, Georgia

INTRODUCTION

Vaccination equity is a national priority, but the extent to which participants in vaccine safety studies reflect the U.S. population is unclear. The Vaccine Safety Datalink (VSD) is a collaboration between the CDC and 13 healthcare organizations and databases, known as “sites,” with over 15 million members as of 2023. Sites contribute electronic health record data toward large studies of vaccine safety and coverage.¹ The aim of this study was to describe the reporting of race and ethnicity in VSD publications from 2011 to 2022.

METHODS

The VSD has collected race and ethnicity data on children since 2000 and adults since 2016. Two authors (K.K., and J.W.) reviewed internal VSD records and external publication lists to identify peer-reviewed publications with VSD data from 2011 to 2022. Studies were classified by year, objective, number of participants or observations, participating VSD sites, exposure, outcome, and race and ethnicity reporting method. The same authors stratified publications by use of mutually exclusive or nonexclusive categories and assessed if study methods met 1997 Office of Management and Budget (OMB) standards for collecting and reporting race and ethnicity data.² Updated OMB guidelines are forthcoming but not yet available³; VSD data collection practices are consistent with 1997 guidelines. This study was exempt from institutional review due to the dataset being publicly available and de-identified.

The proportion of studies reporting participant race and ethnicity over time was assessed using a linear trend via the Cochran–Armitage test. To assess the representativeness of VSD

Address correspondence to: Joshua T.B. Williams, MD, Wellington Webb Pediatric Clinic, 301 W 6th Avenue, MC 1911, Denver CO 80204. joshua.williams@dhha.org.

CREDIT AUTHOR STATEMENT

Kate E. Kurlandsky: Conceptualization, Methodology, Validation, Investigation, Writing – original draft, Writing – review & editing, Visualization, Project administration. Amy B. Stein: Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. Simon J. Hambidge: Conceptualization, Supervision. Eric S. Weintraub: Supervision. Joshua T.B. Williams: Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing, Supervision.

SUPPLEMENTAL MATERIAL

Supplemental materials associated with this article can be found in the online version at <https://doi.org/10.1016/j.amepre.2023.08.019>.

study participants, odds ratios of inclusion (ORs) with 95% CIs were calculated directly using the cross-product ratio of the proportion of participants in racial or ethnic groups versus expected proportions for 2010, 2016, and 2020, using VSD and U.S. census data. Analyses used SAS version 9.4 (SAS Institute Inc., Cary, NC).

RESULTS

There were 224 published articles identified; 215 used VSD data, with 123 (57.2%) vaccine safety studies, 50 (23.3%) methods studies, 31 (14.4%) vaccination coverage and attitudinal studies, and 14 (6.5%) miscellaneous studies. Three studies had co-primary aims and were counted in two categories. Of the 74 studies reporting race and ethnicity, 70 studies used race and ethnicity reporting methods that allowed the total number of participants in each race and ethnicity category to be counted. Sixty articles presented race and ethnicity data on 40,086,769 individuals with mutually exclusive methods, and 10 articles presented race and ethnicity data on 9,047,408 individuals with nonexclusive methods. Thirty-six (48.6%) used 6 mutually exclusive categories: Asian, African American or Black, Hispanic, Other, White, Unknown. Three studies (4.1%) used categories meeting current federal reporting minimums for race and ethnicity from the 1997 Office of Management and Budget guidelines. Overall, 141 studies (65.6%) did not report race or ethnicity. Reporting increased over time ($p < 0.001$); more than 75% of studies reported race and ethnicity by 2022 (Appendix Figure 1). From 2010 to 2020, the odds of study inclusion were lowest for Black VSD members, regardless of reporting method (Table 1).

DISCUSSION

In this retrospective analysis, VSD publications infrequently and heterogeneously reported race and ethnicity data, although reporting increased over time. When participants of racial and ethnic minority groups were compared to populations in 2010–20 VSD and U.S. data, Black individuals were under-represented while Hispanic individuals were over-represented.

Race and ethnicity are sociocultural constructs,⁴ and reporting of race and ethnicity in biomedical research can reinforce incorrect assumptions about racial groups being biologically distinct and decrease attention to sociocultural determinants of health.⁵ Furthermore, existing 1997 OMB standards are outdated and undergoing revisions.³ Yet, inconsistent reporting and variable representation in VSD studies may pose interpretative difficulties for diverse Americans who seek to see themselves represented in vaccine safety research. This inconsistency has been described elsewhere in pediatric clinical trials,⁶ health surveillance,⁷ and hospital system data.⁸

Limitations

Participants may have been counted multiple times, as VSD studies can sample from the same population for separate studies. Many studies did not report race or ethnicity data; unknown data may not be distributed equally.

CONCLUSIONS

VSD publications infrequently and heterogeneously reported race and ethnicity data from 2011 to 2022. Black individuals were under-represented compared to their populations in VSD and U.S. census data. Consistent reporting guidelines and timely updates to federal standards may decrease reporting heterogeneity and increase participant representativeness.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors would like to thank Amelia Jazwa and Tat'Yana Koenigsberg for their contributions to this work. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

This study was supported by the Centers for Disease Control and Prevention (CDC), contract number 75D30122D15426. This activity was reviewed by the CDC and was conducted consistent with applicable federal law and CDC policy. See, for example, 45 C.F.R. part 46.102(l)(2), 21 C.F.R. part 56; 42 U.S. C. §241(d); 5 U.S.C. §552a; 44 U.S.C. §3501 et seq.

The study sponsor, CDC, participated as a coinvestigator and contributed to protocol development, conduct of the study, interpretation of the data, review and revision of the manuscript, approval of the manuscript through official CDC scientific clearance processes, and the decision to submit the manuscript for publication. CDC authors must receive approval through the CDC scientific clearance process to submit an article for publication. The final decision to submit rested with the senior author. The study sponsor did not have the right to direct the submission to a particular journal.

No financial disclosures have been reported by the authors of this paper.

An early version of this work was presented as a poster at the 2022 American Public Health Association Annual Meeting.

REFERENCES

1. McNeil MM, Gee J, Weintraub ES, et al. The Vaccine Safety Datalink: successes and challenges monitoring vaccine safety. *Vaccine*. 2014;32 (42):5390–5398. 10.1016/j.vaccine.2014.07.073. [PubMed: 25108215]
2. Office of Management and Budget. Revision to the standards for the Classification of Federal Data on Race and Ethnicity. Published online 1997 https://obamawhitehouse.archives.gov/omb/fedreg_1997standards. Accessed March 7, 2023.
3. Orvis K Initial proposals for revising the Federal Race and Ethnicity Standards. The White House: Office of Management and Budget. Published online 2023. <https://www.whitehouse.gov/omb/briefing-room/2023/01/26/initial-proposals-for-revising-the-federal-race-and-ethnicity-standards/> Accessed July 24, 2023.
4. Flanagan A, Frey T, Christiansen SL, Bauchner H. The reporting of race and ethnicity in medical and science journals: comments invited. *JAMA*. 2021;325(11):1049. 10.1001/jama.2021.2104. [PubMed: 33616604]
5. Epstein S *Inclusion: The Politics of Difference in Medical Research*. Chicago, IL: The University of Chicago Press, 2007.
6. Rees CA, Stewart AM, Mehta S, et al. Reporting of participant race and ethnicity in published US Pediatric Clinical Trials from 2011 to 2020. *JAMA Pediatr*. 2022;176(5):e220142. 10.1001/jamapediatrics.2022.0142. [PubMed: 35311946]
7. Rodriguez-Lainz A, McDonald M, Fonseca-Ford M, et al. Collection of data on race, ethnicity, language, and nativity by US Public Health Surveillance and Monitoring Systems: gaps and

opportunities. *Public Health Rep.* 2018;133(1):45–54. 10.1177/0033354917745503. [PubMed: 29262290]

8. Cowden JD, Flores G, Chow T, et al. Variability in collection and use of race/ethnicity and language data in 93 pediatric hospitals. *J Racial Ethn Health Disparities.* 2020;7(5):928–936. 10.1007/s40615-020-00716-8. [PubMed: 32056162]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1. Odds of Inclusion^{a,b} by Race and Ethnicity Reporting Category in VSD Studies, Based on VSD and U.S. Census Data, 2011–2022

Reporting method, year(s), and categories	VSD studies			VSD population			U.S. Census		
	Participants	Proportion	Population	Population	Proportion	Odds ratio (95% CI)	Population	Proportion	Odds ratio (95% CI)
Mutually exclusive, 2010									
Black, non-Hispanic	2,910,429	7.26%	622,765	8.15%	0.883 (0.880, 0.885)	37,685,848	12.21%	0.563 (0.562, 0.564)	
Hispanic	11,565,748	28.85%	1,888,393	24.71%	1.236 (1.234, 1.238)	50,477,594	16.35%	2.075 (2.073, 2.076)	
Other ^c , non-Hispanic	7,292,642	18.19%	1,212,654	15.87%	1.179 (1.177, 1.182)	23,764,544	7.70%	2.667 (2.664, 2.669)	
White, non-Hispanic	18,317,950	45.70%	3,919,468	51.28%	0.799 (0.798, 0.801)	196,817,552	63.75%	0.479 (0.478, 0.479)	
Mutually exclusive, 2020									
Black, non-Hispanic	2,910,429	7.26%	753,830	7.42%	0.977 (0.974, 0.979)	39,940,338	12.05%	0.571 (0.571, 0.572)	
Hispanic	11,565,748	28.85%	2,898,889	28.54%	1.016 (1.014, 1.017)	62,080,044	18.73%	1.760 (1.758, 1.761)	
Other ^c , non-Hispanic	7,292,642	18.19%	1,906,079	18.76%	0.963 (0.961, 0.965)	37,731,252	11.38%	1.731 (1.730, 1.733)	
White, non-Hispanic	18,317,950	45.70%	4,600,145	45.28%	1.017 (1.015, 1.018)	191,697,647	57.84%	0.613 (0.613, 0.614)	
Mark all that apply, 2016 (VSD population) ^d and 2010 (U.S. Census)									
Black	1,548,511	8.56%	869,168	7.59%	1.139 (1.136, 1.142)	38,929,319	12.61%	0.649 (0.648, 0.650)	
Hispanic	6,414,870	35.45%	2,785,260	24.33%	1.708 (1.705, 1.711)	50,477,594	16.35%	2.810 (2.807, 2.813)	
Other ^c	4,521,786	24.99%	1,953,100	17.06%	1.620 (1.617, 1.623)	46,262,954	14.98%	1.890 (1.888, 1.892)	
White	11,296,097	62.43%	5,534,173	48.34%	1.775 (1.773, 1.778)	223,553,265	72.41%	0.633 (0.633, 0.635)	
Mark all that apply, 2020									
Black	1,548,511	8.56%	935,881	7.43%	1.165 (1.162, 1.168)	41,104,200	12.40%	0.661 (0.660, 0.662)	
Hispanic	6,414,870	35.45%	3,198,711	25.41%	1.612 (1.610, 1.615)	62,080,044	18.73%	2.383 (2.381, 2.386)	
Other ^c	4,521,786	24.99%	2,280,866	18.12%	1.506 (1.503, 1.508)	86,067,808	25.97%	0.950 (0.949, 0.951)	
White	11,296,097	62.43%	5,765,182	45.80%	1.966 (1.964, 1.969)	204,277,273	61.63%	1.034 (1.033, 1.035)	

^a Analyses were limited to 4 categories due to limitations in published studies' reporting methods, which prevented more granularity.

^b VSD studies used either mutually exclusive race and ethnicity categories or nonexclusive categories (i.e., "mark all that apply"). Table 1 summarizes data by publication type and calculates odds ratios for VSD and U.S. census data, stratified accordingly.

^c "Other": American Indian & Alaskan Native, Asian, Native Hawaiian or Pacific Islander, "Other", and "Multiracial" individuals.

^d Due to changes in race and ethnicity data file coding practices, 2016 is the earliest that "mark all that apply" data were available. VSD, Vaccine Safety DataLink.