


Race and Ethnicity in Real-World Data Sources: Considerations for Medical Device Regulatory Efforts

Journal of Primary Care & Community Health
Volume 12: 1–4
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2150132721994040
journals.sagepub.com/home/jpc


Michelle E. Tarver, MD, PhD¹ 

Abstract

Real-world evidence is increasingly being collected in our interconnected world. This information can provide insights on how medical products are impacting the daily lives of patients, translating into public health consequences. Understanding the demographic characteristics of patients contributing health information to real-world data sources is critical to determining the safety and effectiveness of medical products for specific populations. Race and ethnicity are recorded inconsistently in different real-world data sources, with many lacking reliable information on these variables. To better leverage all forms of real-world data to inform medical device evaluation, accurate and consistent collection of information from different racial and ethnic subgroups is necessary.

Keywords

race, ethnicity, real-world data, regulatory, medical device, electronic health records, registries

Dates received 31 December 2020; revised 19 January 2021; accepted 21 January 2021.

Patient-centered healthcare is proposed as a way to help improve individual and population health outcomes by ensuring that the perspectives of patient are reflected in the care paradigm as well as the systems that collect and record their care journey.¹ To provide this whole-person care, it is critical to understand the sociocultural factors such as race and ethnicity that impact health. In addition, understanding the impact of the medical devices used to provide this care is critical. Medical devices used in the diagnosis, management, and treatment of medical conditions are evaluated by the Food and Drug Administration (FDA) for their safety and effectiveness. Typically, randomized, controlled, clinical trials are the “gold standard” for the evaluation of some medical devices with their selected populations and idealized, controlled conditions. However, clinical trials have been plagued with underrepresentation of diverse racial and ethnic groups. African-derived and Latinx people represent 30% of the US population but only account for 6% of all participants in federally funded clinical trials.² In addition, data shows that participation of racial and ethnic minorities in trials is exponentially lower than the incidence of disease in these groups, with little improvement over time.³ While clinical trials are critical to understanding the performance of some medical devices, they may not always include diverse populations or address the practice challenges encountered daily in the

delivery of healthcare. The inability to account for patient diversity may impact the generalizability of the results to the US public.⁴ Therefore, other sources of information may be useful to help inform public health decisions and patient care.

In the delivery of healthcare, data on clinical, economic, and patient-reported outcomes are generated, collected, curated, and analyzed. Increasingly, health care providers, payers, patients and health delivery systems have demonstrated an interest in using this data to inform clinical care. This data, called real-world data (RWD), is an attractive option to understand the patients’ daily lived experiences with their conditions. As defined in the FDA guidance, RWD are data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources.⁵ Examples of RWD include product and disease registries, claims and billing data, patient-generated data, health surveys, electronic health records (EHRs), and

¹Center for Devices and Radiological Health, Food and Drug Administration, Silver Spring, MD, USA

Corresponding Author:

Michelle E. Tarver, MD, PhD, Center for Devices and Radiological Health, Food and Drug Administration, 10903 New Hampshire Ave, Building 66 Room 5608, Silver Spring, MD 20993, USA.
Email: michelle.tarver@fda.hhs.gov



medical chart reviews.^{5,6} Observational data collected in these data repositories are increasingly being used to guide clinical decision making, determine patient outcomes, and evaluate care paradigms.⁷ RWD may become valid scientific evidence for regulatory decision-making depending on the characteristics and quality of the data. RWD may help bridge the evidentiary gap between research and clinical practice while enhancing efficiency in generating evidence to improve health outcomes.

The 21st Century Cures Act emphasizes that FDA consider how best to use evidence from RWD for reasonable assurance of device safety and effectiveness while accelerating access to important new technologies.⁸ The Center for Devices and Radiological Health (CDRH) is committed to supporting the creation and analysis of RWD for medical device evaluation and surveillance. CDRH has evaluated RWD submitted as evidence in support of expanding the labeled indications of the population for which a device was previously used. For example, the initial approval of transcatheter aortic-valve replacement (TAVR) devices indicated that the procedure was to be done through the femoral artery or through a small incision in the chest. Using national registry data, the FDA expanded the labeled indication to include other routes of device placement. By not requiring a new clinical trial, the device was available to more patients in a shorter time frame.⁹

The full benefit of RWD can only be realized if the necessary data is collected in EHRs, registries, and claims datasets. Many efforts are underway to improve the seamless integration and application of the different types of RWD in regulatory decision-making, including increased use of patient-reported outcomes as part of routine clinical care, widespread use of medical device surveillance initiatives such as the Unique Device Identifier (UDI), and technological advances such as natural language processing. CDRH is collaborating with the Medical Device Innovation Consortium (MDIC), a public-private partnership, in building the National Evaluation System for Health Technology (NEST) to catalyze the timely, reliable, and cost-effective development of evidence using RWD to enhance regulatory and clinical decision-making.¹⁰ NEST creates strategic partnerships and linkages among data sources and the entities that manage them, including registries, electronic health records, payer claims data, patient-generated data, and other sources.¹¹ These data sources historically suffer from inconsistent and incomplete data collection.¹² While still early in development, NEST and other systems like it will require the use of appropriate data quality and methods standards such as ensuring that the identifying data for patients is consistently collected in all the data sources.¹³ It is critical that these data sources contain demographic data on race and ethnicity collected consistently and in a standardized manner to not only facilitate linkages to other data sources, but also to provide insights on device performance across

different demographic groups. To assure that the data being collected can be most useful in informing health-related decisions, it is important that it reflect the composition of the patient population living with the condition.

In 1997, The Office of Management and Budget (OMB) issued revised recommendations for the collection and use of race and ethnicity data by US federal agencies (Policy Directive 15).¹⁴ The recommendation requires that respondents first be asked about ethnicity (i.e., “Hispanic or Latino”) and then asked to identify themselves by race categorized as follows: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White. Not only is there the option to select one or more racial designations, but respondents may also be presented with more granular ethnicity (e.g., Mexican, Puerto Rican) and racial categories (e.g., Chinese, Filipino). The OMB policy also recommends that the information be provided by self-report to minimize misclassification.

The FDA recommends that clinical trials for FDA-regulated medical products use standard terminology for age, sex, gender, race, and ethnicity to ensure that subpopulation data is collected consistently.¹⁵ In addition, the FDA recommends sponsors enroll participants who reflect the demographics of the clinically relevant populations with regard to age, gender, race, and ethnicity.¹⁵ Similar to clinical trials, demographic data recorded in RWD sources are most robust when collected in a consistent and standardized manner. This information may not only help organizations proactively identify and address health disparities, but it also could be used to improve overall public health for all groups. Laws and Heckscher¹⁶ observed that 11.8% of hospital systems studied in New England did not collect any race or ethnicity data. In addition, many systems that collected the information did not conform to the OMB standards. Another survey of US hospitals found that 22% of facilities did not collection information on race and ethnicity.¹⁷ Some investigators have explored approaches to improve the reporting of these important population characteristics. Bhalla et al¹⁸ reported a system-wide, standardized effort to improve the collection of demographic characteristics from a New York-based hospital system. Training registration staff to ask patients to report their demographic information (not registrar observations) dropped the unknown race and ethnicity percentage (47.2% and 62.1%, respectively) to 21.3% and 9.7%, respectively.

In the absence of training patient intake staff or having clear data collection tools, it is common to see misclassification of race and ethnicity in RWD sources. For example, a study examining the correlation between race and ethnicity measures collected in a cancer registry compared to the EHR found that “Hispanic” was often recorded in the race field or that all patients recorded as being “Hispanic” were automatically coded to “white” race.¹⁹ In addition, they

found that the “unknown” or missing race and ethnicity fields resulted in data quality that was worse than if the data had been generated at random.²⁰

Misclassification of race and ethnicity is also a concern for other sources of RWD such as registries. The absence of standardized operational procedures to capture race and ethnicity can plague the quality of data obtained from registries. An analysis of the Surveillance, Epidemiology, and End Results (SEER) registry showed that SEER data underreported the number of cancer patients in specific demographic groups compared to self-identification, except for the white group which was similar. This under-reporting was most marked in American Indians/Alaska Natives which were mostly misclassified as white.²¹ Of note, the demographic data in SEER is subject to the EHR data which may be based on hospital personnel’s subjective appraisal of race and ethnicity instead of self-report.²² Lee et al¹⁹ also evaluated cancer registries and EHRs, finding that there was significant discordance in patient racial and ethnicity descriptions amongst the different health databases. These studies not only highlight the importance of training staff to avoid imputing race and ethnicity, but also the importance of encouraging patients to self-report their demographic characteristics.

Like misclassification, missing data can impede compilation of various data sources as well as valid inferences being drawn from the data. In a comparison of three registries, Mendelsohn et al²³ found an 18% missingness rate for race/ethnicity in one registry. One potential suggestion for improving the data capture for race and ethnicity is a forced-choice response field programmed during the development of the electronic data capture system. However, there are other potential barriers to collecting this information including concerns about patient privacy, potential resistance from patients, and the state laws around soliciting this information. Baker et al²⁰ reported that 17.2% of the Californian’s surveyed were uncomfortable reporting their own race/ethnicity and 46.3% of respondents were worried that providing information could be used to discriminate against them. These findings emphasize the value of including patients in the development of race and ethnicity data collection tools.

By accurately collecting information on race and ethnicity, healthcare systems will enable better understanding of the needs of the populations that they serve, identification of health disparities within their population, and programmatic efforts to improve quality of care, ultimately leading to patient-centered care. Adopting common terminology and standardized approaches to collecting race and ethnicity demographic data such as that described by OMB and FDA will help healthcare systems foster high quality evidence-generation systems that support an inclusive learning health system that benefits society more broadly. The increased adoption of RWD, improved interoperability, and greater patient involvement in evidence generation will

concomitantly promote the use of RWD to help inform regulatory decision-making. We have the opportunity to include the perspectives of patients as we shape the methods used to collect information on their sociocultural experiences through race and ethnicity as well as the outcomes they experience in the healthcare system. By working collaboratively, we can assure that all patients are represented in these rich sources of information and we collectively have the evidence needed to make informed healthcare decisions.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Michelle E. Tarver  <https://orcid.org/0000-0001-8768-6532>

References

1. Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. National Academy Press; 2001. Accessed December 21, 2020. <http://www.jopm.org/evidence/reviews/2011/06/21/spock-feminists-and-the-fight-for-participatory-medicine-a-history/>
2. Oh SS, Galanter J, Thakur N, et al. Diversity in clinical and biomedical research: a promise yet to be fulfilled. *PLoS Med*. 2015;12:e1001918.
3. Loree JM, Anand S, Dasari A, et al. Disparity of race reporting and representation in clinical trials leading to cancer drug approvals from 2008 to 2018. *JAMA Oncol*. 2019;5:e191870.
4. Hohmann AA, Parron DL. How the new NIH guidelines on inclusion of women and minorities apply: efficacy trials, effectiveness trials, and validity. *J Consult Clin Psychol*. 199; 64:851.
5. US Department of Health and Human Services, Food and Drug Administration. Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices. August 31, 2017. Accessed December 31, 2020. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-real-world-evidence-support-regulatory-decision-making-medical-devices>
6. Garrison LP, Neumann PJ, Erickson P, et al. Using real-world data for coverage and payment decisions: The ISPOR real-world data task force report. *Value Health*. 2007;10:326-355.
7. Blumenthal D, Tavenner M. The ‘meaningful use’ regulation for electronic health records. *NEJM*. 2010;363:501-504.
8. Puplic Law. 114–255. December 13, 2016. Accessed December 31, 2020. <https://www.congress.gov/114/plaws/publ255/PLAW-114publ255.pdf>
9. US Department of Health and Human Services, Food and Drug Administration. FDA Facts: Postmarket Patient Registry Ensures Access to Safe and Effective Device. December 21,

2017. Accessed December 31, 2020. <https://www.fda.gov/about-fda/innovation-fda/fda-facts-postmarket-patient-registry-ensures-access-safe-and-effective-devices>
10. NEST Coordinating Center. About Us - NESTcc. Accessed December 21, 2020. <https://nestcc.org/about/about-us>
 11. Shuren J, Califf RM. Need for a national evaluation system for health technology. *JAMA*. 2016;316:1153-1154.
 12. Resnic FS, Matheny ME. Medical devices in the real world. *NEJM*. 2018;378:595-597.
 13. Fleurence RL, Blake K, Shuren J. The future of registries in the era of real-world evidence for medical devices. *JAMA Cardiol*. 2019;4:197-198
 14. Office of Management and Budget (US). Revision to the standards for classification of federated data on race and ethnicity. *Fed Regist*. 1997;62:58781-58790.
 15. US. Department of Health and Human Services, Food and Drug Administration. Collection of Race and Ethnicity Data in Clinical Trials. October 26, 2016. Accessed December 20, 2020. <https://www.fda.gov/media/75453/download>
 16. Laws MB, Heckscher RA. Racial and ethnic identification practices in public health data systems in New England. *Public Health Rep*. 2002;117:50-61.
 17. Hasnain-Wynia R, Pierce D, Pittman MA. Who, when, and how: the current state of race, ethnicity, and primary language data collection in hospitals. Commonwealth Fund Publication No. 726. Health Research and Educational Trust. Accessed December 20, 2020. <https://www.commonwealthfund.org/publications/fund-reports/2004/may/who-when-and-how-current-state-race-ethnicity-and-primary>
 18. Bhalla R, Yongue BG, Currie BP. Standardizing race, ethnicity, and preferred language data collection in hospital information systems: results and implications for healthcare deliver and policy. *J Healthc Qual*. 2011;34:44-52.
 19. Lee SJC, Grobe JE, Tiro JA. Assessing race and ethnicity data quality across cancer registries and EMRs in two hospitals. *J Am Med Inform Assoc*. 2016;23:627-634.
 20. Baker DW, Hasnain-Wynia R, Kandula NR, et al. Attitudes toward health care providers, collecting information about patients' race, ethnicity, and language. *Med Care*. 2007;45:1034-1042.
 21. Clegg LX, Reichman ME, Hankey BF, et al. Quality of race, hispanic ethnicity, and immigrant status in population-based cancer registry data: implications for health disparity studies. *Cancer Causes Control*. 2007;18:177-187.
 22. Gomez SL, Le GM, West DW, et al. Hospital policy and practice regarding the collection of data on race, ethnicity, and birthplace. *Am J Public Health*. 2003;93:1685-1688.
 23. Mendelsohn AB, Dreyer NA, Mattox PW, et al. Characterization of missing data in clinical registry studies. *Ther Innov Regul Sci*. 2015;49:146-154.