

Original Paper

# Information Loss in Harmonizing Granular Race and Ethnicity Data: Descriptive Study of Standards

Karen Wang<sup>1,2</sup>, MD, MS; Holly Grossetta Nardini<sup>3</sup>, MLS; Lori Post<sup>4</sup>, PhD; Todd Edwards<sup>5</sup>, MS, PhD; Marcella Nunez-Smith<sup>1</sup>, MD, MHS; Cynthia Brandt<sup>2,6</sup>, MD, MPH

<sup>1</sup>Equity Research and Innovation Center, General Internal Medicine, Yale School of Medicine, New Haven, CT, United States

<sup>2</sup>Center for Medical Informatics, Yale School of Medicine, New Haven, CT, United States

<sup>3</sup>Harvey Cushing/John Hay Whitney Medical Library, Yale School of Medicine, New Haven, CT, United States

<sup>4</sup>Buehler Center for Health Policy and Economics, Feinberg School of Medicine, Chicago, IL, United States

<sup>5</sup>Division of Epidemiology, Department of Medicine, Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, United States

<sup>6</sup>Veteran Affairs Connecticut Healthcare System, US Department of Veteran Affairs, West Haven, CT, United States

**Corresponding Author:**

Karen Wang, MD, MS

Equity Research and Innovation Center

General Internal Medicine

Yale School of Medicine

100 Church Street South, A200

New Haven, CT, 06520

United States

Phone: 1 203 785 5233

Email: [karen.wang@yale.edu](mailto:karen.wang@yale.edu)

## Abstract

**Background:** Data standards for race and ethnicity have significant implications for health equity research.

**Objective:** We aim to describe a challenge encountered when working with a multiple-race and ethnicity assessment in the Eastern Caribbean Health Outcomes Research Network (ECHORN), a research collaborative of Barbados, Puerto Rico, Trinidad and Tobago, and the US Virgin Islands.

**Methods:** We examined the data standards guiding harmonization of race and ethnicity data for multiracial and multiethnic populations, using the Office of Management and Budget (OMB) Statistical Policy Directive No. 15.

**Results:** Of 1211 participants in the ECHORN cohort study, 901 (74.40%) selected 1 racial category. Of those that selected 1 category, 13.0% (117/901) selected Caribbean; 6.4% (58/901), Puerto Rican or Boricua; and 13.5% (122/901), the mixed or multiracial category. A total of 17.84% (216/1211) of participants selected 2 or more categories, with 15.19% (184/1211) selecting 2 categories and 2.64% (32/1211) selecting 3 or more categories. With aggregation of ECHORN data into OMB categories, 27.91% (338/1211) of the participants can be placed in the “more than one race” category.

**Conclusions:** This analysis exposes the fundamental informatics challenges that current race and ethnicity data standards present to meaningful collection, organization, and dissemination of granular data about subgroup populations in diverse and marginalized communities. Current standards should reflect the science of measuring race and ethnicity and the need for multidisciplinary teams to improve evolving standards throughout the data life cycle.

(*J Med Internet Res* 2020;22(7):e14591) doi: [10.2196/14591](https://doi.org/10.2196/14591)

**KEYWORDS**

continental population groups; multiracial populations; multiethnic groups; data standards; health status disparities; race factors; demography

## Introduction

The cutting edge of precision medicine, such as the integration of contextual data about the social determinants of health with individual health data and the leveraging of data from across different studies, is seen as a mechanism to innovate and solve health problems for all populations [1,2]. These solutions can only be realized by grounding them in the concept of health equity [3,4]. According to the Robert Wood Johnson Foundation [5],

*Health equity means that everyone has a fair and just opportunity to be healthier. This requires removing obstacles to health such as poverty, discrimination, and their consequences, including powerlessness and lack of access to good jobs with fair pay, quality education and housing, safe environments, and health care.*

To attain health equity, the research community must identify where health disparities exist but may inadvertently exacerbate health disparities by failing to identify invisible or at-risk populations. Without this information, we may reinforce inequities rather than identify policies, laws, systems, environments, and practices that can improve opportunities for health in these communities [6-9].

There is a significant body of literature about the collection of and proposed definitions for more granular data for race, ethnicity, and other demographic data in the health sciences that would better identify at-risk populations. In some areas, more granular data collection is used to be more inclusive of the ways that individuals self-identify and to examine how racism and discrimination affect health in these populations [10-13]. However, the research community has not sufficiently considered how to meaningfully organize granular data about subgroup populations in diverse communities so that these data can be used across multiple studies to address the challenges in these communities [5,14]. In this paper, we used our experience with a cohort study in the eastern Caribbean to illustrate the need for an updated and comprehensive data standard to enable future data integration and sharing of data sources for diverse multiracial populations.

## Methods

The Eastern Caribbean Health Outcomes Research Network (ECHORN) cohort study follows community-dwelling adults older than 40 years residing in Barbados, Trinidad, Tobago, and the United States' territories of the US Virgin Islands and Puerto Rico [15]. Baseline participants were enrolled between 2013 and 2016 and completed a questionnaire to capture self-reported sociodemographics and health-related information. In the questionnaire, participants had the option to self-identify race and ethnicity by selecting any of the items listed: mixed or multiracial, white, black or African, Caribbean, Asian, East Indian, Hispanic or Latino, Puerto Rican or Boricua, other, or prefer not to answer. This list was developed with stakeholder input to reflect the current scientific literature on the measurement of race and ethnicity in epidemiologic and health

outcomes research. Excluding the choice of "other", there were 256 potential answers.

The National Institutes of Health (NIH) requires investigators to collect and report race and ethnicity based on Office of Management and Budget (OMB) Statistical Policy Directive No. 15 data standards, developed in 1977 and updated in 2007 [16]. The standards include 7 race and ethnicity categories, with instructions to use 2 questions. The first question collects ethnicity information (Hispanic/Latino or not Hispanic/Latino), and the second provides the option to select more than one of the 5 racial categories (American Indian or Alaska Native, Asian, black or African American, Native Hawaiian or other Pacific Islander, or white). Data reported are required to include (1) the number of respondents in each ethnic category, (2) the number of respondents who selected only 1 of 5 racial categories, (3) the number of respondents who selected multiple racial categories, and (4) the number of respondents in each racial category who identified as Hispanic or Latino. Reporting detailed distributions of multiple responses should be aggregated into the required categories.

## Results

In an analysis of ECHORN data from 1211 participants, 901 (74.40%) selected 1 racial category. Of those that selected only 1 category, 13.0% (117/901) selected Caribbean; 6.4% (58/901), Puerto Rican or Boricua; and 13.5% (122/901), mixed or multiracial. A total of 17.84% (216/1211) of participants selected 2 or more categories, with 15.19% (184/1211) selecting 2 categories and 2.64% (32/1211) selecting 3 or more categories. The participants who selected more than one category included but were not limited to white Hispanic, black Hispanic, black Asian, black East Indian, and Caribbean Hispanic individuals. Based on OMB data standards, by collapsing those who selected multiple groups and those who selected only the multiracial and mixed category, 27.91% (338/1211) of the ECHORN participants are placed in the "more than one race" category. The nuanced race and ethnicity data become simply "more than one race, black or African American, white, Hispanic or Latino, Asian, or not reported/unknown."

## Discussion

### Principal Findings

Data standards for race and ethnicity need to reflect and evolve with the scientific advances around the measurement and mechanisms by which race and ethnicity can affect health [17,18]. The stated goals of the OMB standards are to harmonize the collection and presentation of population race and ethnicity information across federal databases and to facilitate comparisons and analyses [16]. To achieve the objectives of precision medicine, ECHORN data on those who selected multiple categories, such as the white Hispanic, black Hispanic, black Asian, black East Indian, or Caribbean Hispanic individuals, could be aggregated with other data sources with defined granular race and ethnicity data on this population. For example, the US Census Bureau has developed granular data collection standards on race and ethnicity, quantifying the growth of these multiracial and multiethnic populations. The

State of Massachusetts, along with other states, has for several decades mandated the collection of more granular race and ethnicity information, with over 30 categories on many data collection forms, including birth certificates and clinical data [19-21]. However, it is not specified how multiracial and multiethnic populations' granular data can be harmonized to facilitate comparisons and analyses across any datasets collected using different standards. Current research collaborations, such as Observational Health Data Sciences and Informatics, that are focused on facilitating harmonization and sharing data by using standard ontologies such as Systematized Nomenclature of Medicine - Clinical Terms and Logical Observation Identifiers Names and Codes, have no agreed-upon standard for granular race and ethnicity data or a mechanism to map across data standards [22].

Standards have not been refined or expanded adequately to accommodate studies collecting granular race and ethnicity data on mixed or multiracial individuals [23,24], and these imprecise standards have unmeasured effects on health. Current standards may contribute to health disparities by providing insights on persons privileged by these data structures and by further marginalizing persons whose racial and ethnic identity is obscured in suboptimal data standards for classification and harmonization. While the OMB standards clearly state that "the racial and ethnic categories set forth in the standards should not be interpreted as being primarily biological or genetic in reference," [16] current standards used in health sciences continue to support the conflation of the social categories of race and ethnicity with the biologic and genetic categories in day-to-day practice. This conflation has powerful implications for population-level biomedical discoveries and clinical care and treatment [25]. For example, inherent inequities in data standards are reflected in the continued use of monoracial guidelines as the standard of care in clinical practice, in spite of knowledge that suggests the fallacies of this framework [26,27]. It also has implications for population genetics and the translation of discoveries. These social constructs are used to develop standards for genetic population norms, wherein outlier data are discarded and study findings are associated with a singular racial group, influencing the development of biologic treatments directed at mutations that are predominantly in a particular racial group [28,29]. However, there is no biological origin for much of what we define as health inequities [30]. At the individual level, understanding how contextual factors, such as racism and discrimination, affect how a multiracial and multiethnic person understands and addresses their own health risks is significantly limited because of the lack of data and data

standards that organize and share granular data in a viable way [31]. Importantly, multiracial and multiethnic populations are only one example of diverse subgroups affected by this lack of evolved standards; we risk the loss of rich information on the individuals who select more than one category for their race and ethnicity.

Several incremental changes can be instituted to improve these data standards on race and ethnicity. In addition to standardizing the definitions and collection of more granular race and ethnicity data across research studies, as advocated by numerous prior researchers [10-13], the standards need to be developed and adopted by funding agencies, such as the NIH, so that researchers from different groups who are collecting granular data can share and aggregate data [32,33]. These comprehensive data standards need to specify how to systematically categorize self-identified granular race and ethnicity data for those who have selected 2 or more categories. The standards can provide mechanisms to map race and ethnicity data collection standards to other standards over time, such as the Centers for Disease Control and Prevention Race and Ethnicity Code Set, Version 1.0, or the United States Census Bureau [32-34]. As data are collected and shared between regional and national entities, this mapping between standards will enable the aggregation of smaller data about underrepresented populations across studies and global contexts [32,35,36] and will provide a systematic method for individuals who have collected granular data to organize and map their data to less-evolved data collection standards.

## Conclusion

Comprehensive data standards for race and ethnicity throughout the life course of health science research studies are critical to identifying and achieving health equity for populations differentially affected by discrimination [23,24,37]. Partnerships with relevant stakeholders in the development and mapping of standards are essential and will likely increase availability of meaningful and usable data [38]. Multidisciplinary research teams that focus on health equity (eg, informaticists, data scientists, information scientists, geneticists, sociologists, and public health researchers) need to adapt standards and advance the theory and systems for organizing, storing, and analyzing complex data. These standards are urgently needed to ensure the sustained value of data collected about diverse populations and their health. With the status quo, we will continue to privilege one group's data over another's, lack meaningful and useable information of diverse racial and ethnic population groups, and likely further perpetuate health inequities.

## Acknowledgments

We acknowledge Ms Hannah Friedman, who was a research assistant on this project, and Ms Mary Miller and Ms Natasha Wenner for their editorial assistance.

## Conflicts of Interest

Drs KW, CB, and MNS are supported in part by U54MD010711 from the National Institute on Minority Health and Health Disparities. The funder had no role in the design or conduct of the manuscript. Views expressed are those of the authors and do not represent those of the funding source.

## References

<http://www.jmir.org/2020/7/e14591/>

J Med Internet Res 2020 | vol. 22 | iss. 7 | e14591 | p. 3  
(page number not for citation purposes)

1. Fuchs VR. Social Determinants of Health: Caveats and Nuances. *JAMA* 2017 Jan 03;317(1):25-26. [doi: [10.1001/jama.2016.17335](https://doi.org/10.1001/jama.2016.17335)] [Medline: [28030707](https://pubmed.ncbi.nlm.nih.gov/28030707/)]
2. Breen N, Jackson JS, Wood F, Wong DW, Zhang X. Translational Health Disparities Research in a Data-Rich World. *Am J Public Health* 2019 Jan;109(S1):S41-S42. [doi: [10.2105/AJPH.2019.304951](https://doi.org/10.2105/AJPH.2019.304951)] [Medline: [30699034](https://pubmed.ncbi.nlm.nih.gov/30699034/)]
3. Schinasi LH, Auchincloss AH, Forrest CB, Diez Roux AV. Using electronic health record data for environmental and place based population health research: a systematic review. *Ann Epidemiol* 2018 Jul;28(7):493-502. [doi: [10.1016/j.annepidem.2018.03.008](https://doi.org/10.1016/j.annepidem.2018.03.008)] [Medline: [29628285](https://pubmed.ncbi.nlm.nih.gov/29628285/)]
4. Braveman P. Health disparities and health equity: concepts and measurement. *Annu Rev Public Health* 2006;27:167-194. [doi: [10.1146/annurev.publhealth.27.021405.102103](https://doi.org/10.1146/annurev.publhealth.27.021405.102103)] [Medline: [16533114](https://pubmed.ncbi.nlm.nih.gov/16533114/)]
5. Braveman P, Arkin E, Orleans T, Proctor D, Plough A. What is Health Equity? Robert Wood Johnson Foundation. URL: <https://www.rwjf.org/en/library/research/2017/05/what-is-health-equity-.html> [accessed 2020-05-02]
6. Bourgois P, Holmes SM, Sue K, Quesada J. Structural Vulnerability: Operationalizing the Concept to Address Health Disparities in Clinical Care. *Acad Med* 2017 Mar;92(3):299-307 [FREE Full text] [doi: [10.1097/ACM.0000000000001294](https://doi.org/10.1097/ACM.0000000000001294)] [Medline: [27415443](https://pubmed.ncbi.nlm.nih.gov/27415443/)]
7. Khoury MJ, Iademarco MF, Riley WT. Precision Public Health for the Era of Precision Medicine. *Am J Prev Med* 2016 Mar;50(3):398-401 [FREE Full text] [doi: [10.1016/j.amepre.2015.08.031](https://doi.org/10.1016/j.amepre.2015.08.031)] [Medline: [26547538](https://pubmed.ncbi.nlm.nih.gov/26547538/)]
8. Anderson LM, Adeney KL, Shinn C, Safranek S, Buckner-Brown J, Krause LK. Community coalition-driven interventions to reduce health disparities among racial and ethnic minority populations. *Cochrane Database Syst Rev* 2015 Jun 15(6):CD009905. [doi: [10.1002/14651858.CD009905.pub2](https://doi.org/10.1002/14651858.CD009905.pub2)] [Medline: [26075988](https://pubmed.ncbi.nlm.nih.gov/26075988/)]
9. Betancourt JR, Green AR, Carrillo J, Ananeh-Firempong O. Defining cultural competence: a practical framework for addressing racial/ethnic disparities in health and health care. *Public Health Rep* 2003;118(4):293-302 [FREE Full text] [doi: [10.1093/phr/118.4.293](https://doi.org/10.1093/phr/118.4.293)] [Medline: [12815076](https://pubmed.ncbi.nlm.nih.gov/12815076/)]
10. Ford CL, Harawa NT. A new conceptualization of ethnicity for social epidemiologic and health equity research. *Soc Sci Med* 2010 Jul;71(2):251-258 [FREE Full text] [doi: [10.1016/j.socscimed.2010.04.008](https://doi.org/10.1016/j.socscimed.2010.04.008)] [Medline: [20488602](https://pubmed.ncbi.nlm.nih.gov/20488602/)]
11. Griffith DM, Moy E, Reischl TM, Dayton E. National data for monitoring and evaluating racial and ethnic health inequities: where do we go from here? *Health Educ Behav* 2006 Aug;33(4):470-487. [doi: [10.1177/1090198106287923](https://doi.org/10.1177/1090198106287923)] [Medline: [16769756](https://pubmed.ncbi.nlm.nih.gov/16769756/)]
12. Mays VM, Ponce NA, Washington DL, Cochran SD. Classification of race and ethnicity: implications for public health. *Annu Rev Public Health* 2003;24:83-110 [FREE Full text] [doi: [10.1146/annurev.publhealth.24.100901.140927](https://doi.org/10.1146/annurev.publhealth.24.100901.140927)] [Medline: [12668755](https://pubmed.ncbi.nlm.nih.gov/12668755/)]
13. Williams DR, Lavizzo-Mourey R, Warren RC. The concept of race and health status in America. *Public Health Rep* 1994;109(1):26-41 [FREE Full text] [Medline: [8303011](https://pubmed.ncbi.nlm.nih.gov/8303011/)]
14. Sankar PL, Parker LS. The Precision Medicine Initiative's All of Us Research Program: an agenda for research on its ethical, legal, and social issues. *Genet Med* 2017 Jul;19(7):743-750. [doi: [10.1038/gim.2016.183](https://doi.org/10.1038/gim.2016.183)] [Medline: [27929525](https://pubmed.ncbi.nlm.nih.gov/27929525/)]
15. Wang KH, Thompson TA, Galusha D, Friedman H, Nazario CM, Nunez M, ECHORN Writing Group. Non-communicable chronic diseases and timely breast cancer screening among women of the Eastern Caribbean Health Outcomes Research Network (ECHORN) Cohort Study. *Cancer Causes Control* 2018 Mar;29(3):315-324 [FREE Full text] [doi: [10.1007/s10552-018-1005-4](https://doi.org/10.1007/s10552-018-1005-4)] [Medline: [29423760](https://pubmed.ncbi.nlm.nih.gov/29423760/)]
16. NIH Policy on Reporting Race and Ethnicity Data: Subjects in Clinical Research. National Institutes of Health.: National Institutes of Health; 2001 Aug 08. URL: <https://grants.nih.gov/grants/guide/notice-files/not-od-01-053.html> [accessed 2020-05-07]
17. Goodman AH. Why genes don't count (for racial differences in health). *Am J Public Health* 2000 Nov;90(11):1699-1702. [doi: [10.2105/ajph.90.11.1699](https://doi.org/10.2105/ajph.90.11.1699)] [Medline: [11076233](https://pubmed.ncbi.nlm.nih.gov/11076233/)]
18. Roth WD. The multiple dimensions of race. *Ethnic and Racial Studies* 2016 Mar 21;39(8):1310-1338. [doi: [10.1080/01419870.2016.1140793](https://doi.org/10.1080/01419870.2016.1140793)]
19. Jorgensen S, Thorlby R, Weinick RM, Ayanian JZ. Responses of Massachusetts hospitals to a state mandate to collect race, ethnicity and language data from patients: a qualitative study. *BMC Health Serv Res* 2010 Dec 31;10:352 [FREE Full text] [doi: [10.1186/1472-6963-10-352](https://doi.org/10.1186/1472-6963-10-352)] [Medline: [21194450](https://pubmed.ncbi.nlm.nih.gov/21194450/)]
20. Hawkins SS, Cohen BB. Affordable Care Act standards for race and ethnicity mask disparities in maternal smoking during pregnancy. *Prev Med* 2014 Aug;65:92-95 [FREE Full text] [doi: [10.1016/j.ypmed.2014.05.002](https://doi.org/10.1016/j.ypmed.2014.05.002)] [Medline: [24823903](https://pubmed.ncbi.nlm.nih.gov/24823903/)]
21. Folkerts G, van Esch B, Janssen M, Nijkamp FP. Superoxide production by broncho-alveolar cells is diminished in parainfluenza-3 virus treated guinea pigs. *Agents Actions Suppl* 1990;31:139-142. [doi: [10.1007/978-3-0348-7379-6\\_18](https://doi.org/10.1007/978-3-0348-7379-6_18)] [Medline: [1964360](https://pubmed.ncbi.nlm.nih.gov/1964360/)]
22. Race and Ethnicity in the OMOP CDM [discussion forum]. *Observational Health Sciences Data and Informatics*. URL: <https://forums.ohdsi.org/t/race-and-ethnicity-in-the-omop-cdm/8700> [accessed 2020-05-02]
23. Efthimiadis E, Afifi M. Population groups: indexing, coverage, and retrieval effectiveness of ethnically related health care issues in health sciences databases. *Bull Med Libr Assoc* 1996 Jul;84(3):386-396 [FREE Full text] [Medline: [8883987](https://pubmed.ncbi.nlm.nih.gov/8883987/)]
24. Aspinall PJ. The operationalization of race and ethnicity concepts in medical classification systems: issues of validity and utility. *Health Informatics J* 2016 Jul 25;11(4):259-274. [doi: [10.1177/1460458205055688](https://doi.org/10.1177/1460458205055688)]



25. Kahn J. Genes, race, and population: avoiding a collision of categories. *Am J Public Health* 2006 Nov;96(11):1965-1970. [doi: [10.2105/AJPH.2005.067926](https://doi.org/10.2105/AJPH.2005.067926)] [Medline: [17018831](https://pubmed.ncbi.nlm.nih.gov/17018831/)]
26. Cortes-Bergoderi M, Thomas RJ, Albuquerque FN, Batsis JA, Burdiat G, Perez-Terzic C, et al. Validity of cardiovascular risk prediction models in Latin America and among Hispanics in the United States of America: a systematic review. *Rev Panam Salud Publica* 2012 Aug;32(2):131-139. [doi: [10.1590/s1020-49892012000800007](https://doi.org/10.1590/s1020-49892012000800007)] [Medline: [23099874](https://pubmed.ncbi.nlm.nih.gov/23099874/)]
27. DeFilippis AP, Young R, McEvoy JW, Michos ED, Sandfort V, Kronmal RA, et al. Risk score overestimation: the impact of individual cardiovascular risk factors and preventive therapies on the performance of the American Heart Association-American College of Cardiology-Atherosclerotic Cardiovascular Disease risk score in a modern multi-ethnic cohort. *Eur Heart J* 2017 Feb 21;38(8):598-608 [FREE Full text] [doi: [10.1093/eurheartj/ehw301](https://doi.org/10.1093/eurheartj/ehw301)] [Medline: [27436865](https://pubmed.ncbi.nlm.nih.gov/27436865/)]
28. Jameson JL, Longo DL. Precision medicine--personalized, problematic, and promising. *N Engl J Med* 2015 Jun 04;372(23):2229-2234. [doi: [10.1056/NEJMs1503104](https://doi.org/10.1056/NEJMs1503104)] [Medline: [26014593](https://pubmed.ncbi.nlm.nih.gov/26014593/)]
29. Spratt DE, Chan T, Waldron L, Speers C, Feng FY, Ogunwobi OO, et al. Racial/Ethnic Disparities in Genomic Sequencing. *JAMA Oncol* 2016 Aug 01;2(8):1070-1074 [FREE Full text] [doi: [10.1001/jamaoncol.2016.1854](https://doi.org/10.1001/jamaoncol.2016.1854)] [Medline: [27366979](https://pubmed.ncbi.nlm.nih.gov/27366979/)]
30. Suzuki K, Von Vacano DA, editors. *Reconsidering Race: Social Science Perspectives on Racial Categories in the Age of Genomics*. Oxford, England: Oxford University Press; 2018.
31. Burchard EG, Ziv E, Coyle N, Gomez SL, Tang H, Karter AJ, et al. The importance of race and ethnic background in biomedical research and clinical practice. *N Engl J Med* 2003 Mar 20;348(12):1170-1175. [doi: [10.1056/NEJMs025007](https://doi.org/10.1056/NEJMs025007)] [Medline: [12646676](https://pubmed.ncbi.nlm.nih.gov/12646676/)]
32. Institute of Medicine, Subcommittee on Standardized Collection of Race/Ethnicity Data for Healthcare Quality Improvement. In: Ulmer C, McFadden B, Nerenz DR, editors. *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement*. Washington, DC: National Academies Press; 2009.
33. Strmic-Pawl HV, Jackson BA, Garner S. Race Counts: Racial and Ethnic Data on the U.S. Census and the Implications for Tracking Inequality. *Sociology of Race and Ethnicity* 2017 Dec 04;4(1):1-13. [doi: [10.1177/2332649217742869](https://doi.org/10.1177/2332649217742869)]
34. Race and Ethnicity Code Set Version 1. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/phn/resources/vocabulary/documents/cdc-race--ethnicity-background-and-purpose.pdf> [accessed 2020-05-02]
35. Representing Patient Race and Ethnicity. HealthIT.gov. URL: <https://www.healthit.gov/isa/representing-patient-race-and-ethnicity> [accessed 2020-05-02]
36. Ethnic group, national identity, and religion. Office for National Statistics. URL: <https://www.ons.gov.uk/methodology/classificationsandstandards/measuringequality/ethnicgroupnationalidentityandreligion> [accessed 2020-05-02]
37. Aspinall P. The categorization of African descent populations in Europe and the USA: should lexicons of recommended terminology be evidence-based? *Public Health* 2008 Jan;122(1):61-69. [doi: [10.1016/j.puhe.2007.02.018](https://doi.org/10.1016/j.puhe.2007.02.018)] [Medline: [17645900](https://pubmed.ncbi.nlm.nih.gov/17645900/)]
38. Ward M, Schulz AJ, Israel BA, Rice K, Martenies SE, Markarian E. A conceptual framework for evaluating health equity promotion within community-based participatory research partnerships. *Eval Program Plann* 2018 Oct;70:25-34 [FREE Full text] [doi: [10.1016/j.evalprogplan.2018.04.014](https://doi.org/10.1016/j.evalprogplan.2018.04.014)] [Medline: [29894902](https://pubmed.ncbi.nlm.nih.gov/29894902/)]

## Abbreviations

**ECHORN:** Eastern Caribbean Health Outcomes Research Network

**NIH:** National Institutes of Health

**OMB:** Office of Management and Budget

*Edited by G Eysenbach; submitted 03.05.19; peer-reviewed by P Aspinall, D Griffith; comments to author 29.09.19; revised version received 24.02.20; accepted 12.03.20; published 20.07.20*

*Please cite as:*

Wang K, Grossetta Nardini H, Post L, Edwards T, Nunez-Smith M, Brandt C

*Information Loss in Harmonizing Granular Race and Ethnicity Data: Descriptive Study of Standards*

*J Med Internet Res* 2020;22(7):e14591

URL: <http://www.jmir.org/2020/7/e14591/>

doi: [10.2196/14591](https://doi.org/10.2196/14591)

PMID: [32706693](https://pubmed.ncbi.nlm.nih.gov/32706693/)

©Karen Wang, Holly Grossetta Nardini, Lori Post, Todd Edwards, Marcella Nunez-Smith, Cynthia Brandt. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>), 20.07.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the *Journal of Medical Internet*

Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.