Data justice and community-engagement: the collection and use of race and ethnicity data in the EHR

Karen Wang, MD MHS
Assistant Professor
General Internal Medicine/Health Informatics
Core Faculty, Equity Research and Innovation Center (ERIC)
karen.wang@yale.edu

The National Academies, Health and Medicine Division

January 31, 2024

Yale SCHOOL OF MEDICINE

There is no commercial support for this presentation

There are no personal or financial relationships with commercial interests relevant to this presentation in the past 12 months

Today's Discussion

1. Background: Data justice framework and a community-engaged approach for data collection and use
2. Current issues around collection and use of race and ethnicity data

Known issues in the use of race and ethnicity data

- Poor quality data to examine health disparities leveraging existing technologies
- Use of data that harms racially minoritized communities
- Limited use of these data to address existing health inequities

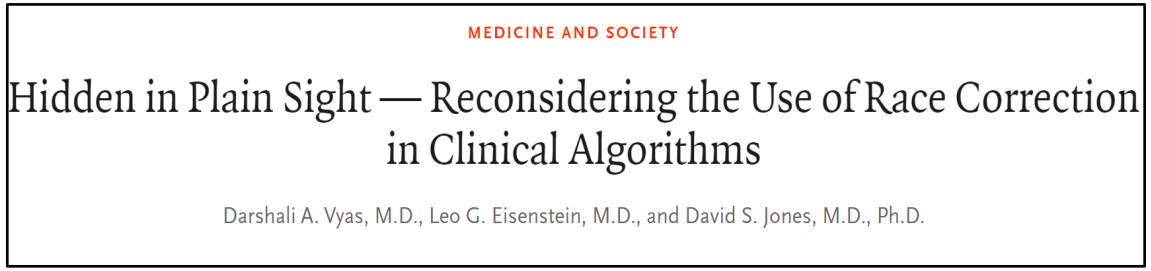
CORRESPONDENCE | VOLUME 397, ISSUE 10281, P1259-1260, APRIL 03, 2021

Missing again: US racial and ethnic data for COVID-19 vaccination

Nancy Krieger • Pamela D Waterman • Jarvis T Chen • Christian Testa • William P Hanage

Published: March 17, 2021 • DOI: https://doi.org/10.1016/S0140-6736(21)00465-7

In the USA, Feb 2 is Groundhog Day, when the famous Punxsutawney, PA, groun
predicts how long winter will last. The phrase also now evokes thoughts of an e
which an egotistical cynical weather reporter who is assigned to film the ground
the day. He escapes only by learning the errors of his ways, redeemed by self-re
authentic love for another person.[1]



MEDICINE AND SOCIETY

Hidden in Plain Sight — Reconsidering the Use of Race Correction in Clinical Algorithms

Darshali A. Vyas, M.D., Leo G. Eisenstein, M.D., and David S. Jones, M.D., Ph.D.

https://www.nejm.org/doi/full/10.1056/NEJMms2004740

Data justice framework is a mechanism to improve meaningful information and use of technologies for health equity

Exploring Data Justice: Conceptions, Applications and Directions

Lina Dencik, Arne Hintz, Joanna Redden & Emiliano Treré

To cite this article: Lina Dencik, Arne Hintz, Joanna Redden & Emiliano Treré (2019) Exploring Data Justice: Conceptions, Applications and Directions, Information, Communication & Society, 22:7, 873-881, DOI: 10.1080/1369118X.2019.1606268

To link to this article: https://doi.org/10.1080/1369118X.2019.1606268

- "engage with data in a way that explicitly asks questions of power, politics, inclusion and interest"
- Engages with trust, accountability governance -- with who decides, owns, uses, benefits

The CARE Principles for Indigenous Data Governance

Research Data Alliance International Indigenous Data Sovereignty Interest Group. (September 2019). "CARE Principles for Indigenous Data Governance." The Global Indigenous Data Alliance https://www.gida-global.org/care, drafted 2018

Carroll, S.R., Rodriguez-Lonebear, D. and Martinez, A., 2019. Indigenous Data Governance: Strategies from United States Native Nations. *Data Science Journal*, 18(1), p.31. DOI: http://doi.org/10.5334/dsj-2019-031

Community-engaged informatics approach

Integrating community-based participatory research and informatics approaches to improve the engagement and health of underserved populations

Kim M Unertl, Chris L Schaefbauer, Terrance R Campbell, Charles Senteio, Katie A Siek, Suzanne Bakken, Tiffany C Veinot

Journal of the American Medical Informatics Association, Volume 23, Issue 1, January 2016, Pages 60–73

Published: 30 Ju

AMIA Annu Symp Proc. 2017; 2017: 1715–1723. PMCID: PMC5977588

Published online 2018 Apr 16. PMID: 29854242

Using a community-engaged health informatics approach to develop a web analytics research platform for sharing data with community stakeholders

Karen H. Wang, MD MHS,[1] Luis Marenco, MD,[1,3] Johanna Elumn Madera, PhD,[2] Jenerius A. Aminawung, MD, MPH,[1] Emily A. Wang, MD MAS,[1] and Kei-Hoi Cheung, PhD[1,3]

Centering Equity In The Design And Use Of Health Information Systems: Partnering With Communities On Race, Ethnicity, And Language Data

Karen Wang, Theresa Cullen, Marcella Nunez-Smith

MAY 19, 2021 10.1377/hblog20210514.126700

- Engagement of communities
  - design and use of health technology for clinical research and care
  - collection and use of their data for clinical research and care
    - race and ethnicity
    - gender identity
  - use of these data and technologies for community benefit

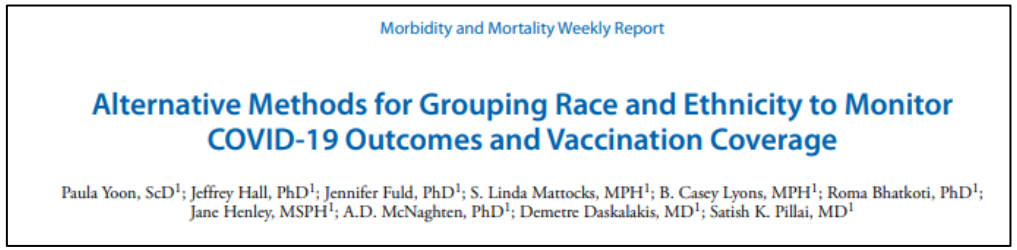
How do we meaningfully use granular data of people who have selected multiple race and ethnicity categories?

Why is this question important to consider

- Number of people selecting more than one category is increasing
  - Since 2000, US Census has had option to select more than one racial category
  - 2010->2020: Increased from ~3% (9 million people) to ~10% of the population (33.8 million people)
- Collection standards change
  - US Census study demonstrates combining the standard separate race and ethnicity questions into one question to improve data quality (NCT 2015)
- Healthcare systems are collecting more granular race and ethnicity data

https://www.pewresearch.org/fact-tank/2020/02/25/the-changing-categories-the-u-s-has-used-to-measure-race/
https://www.pewresearch.org/social-trends/2015/06/11/chapter-1-race-and-multiracial-americans-in-the-u-s-census/
https://www.census.gov/library/stories/2021/08/2020-united-states-population-more-racially-ethnically-diverse-than-2010.html
https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/plan/final-analysis/2015nct-race-ethnicity-analysis.html

Standards to use data from people who have selected multiple race and ethnicity categories are limited



Morbidity and Mortality Weekly Report

Alternative Methods for Grouping Race and Ethnicity to Monitor COVID-19 Outcomes and Vaccination Coverage

Paula Yoon, ScD[1]; Jeffrey Hall, PhD[1]; Jennifer Fuld, PhD[1]; S. Linda Mattocks, MPH[1]; B. Casey Lyons, MPH[1]; Roma Bhatkoti, PhD[1]; Jane Henley, MSPH[1]; A.D. McNaghten, PhD[1]; Demetre Daskalakis, MD[1]; Satish K. Pillai, MD[1]

Method A

Race/Ethnicity groups

- American Indian/Alaska Native, non-Hispanic
- Asian, non-Hispanic
- Black or African American, non-Hispanic
- Hispanic
- Native Hawaiian or Other Pacific Islander, non-Hispanic
- White, non-Hispanic

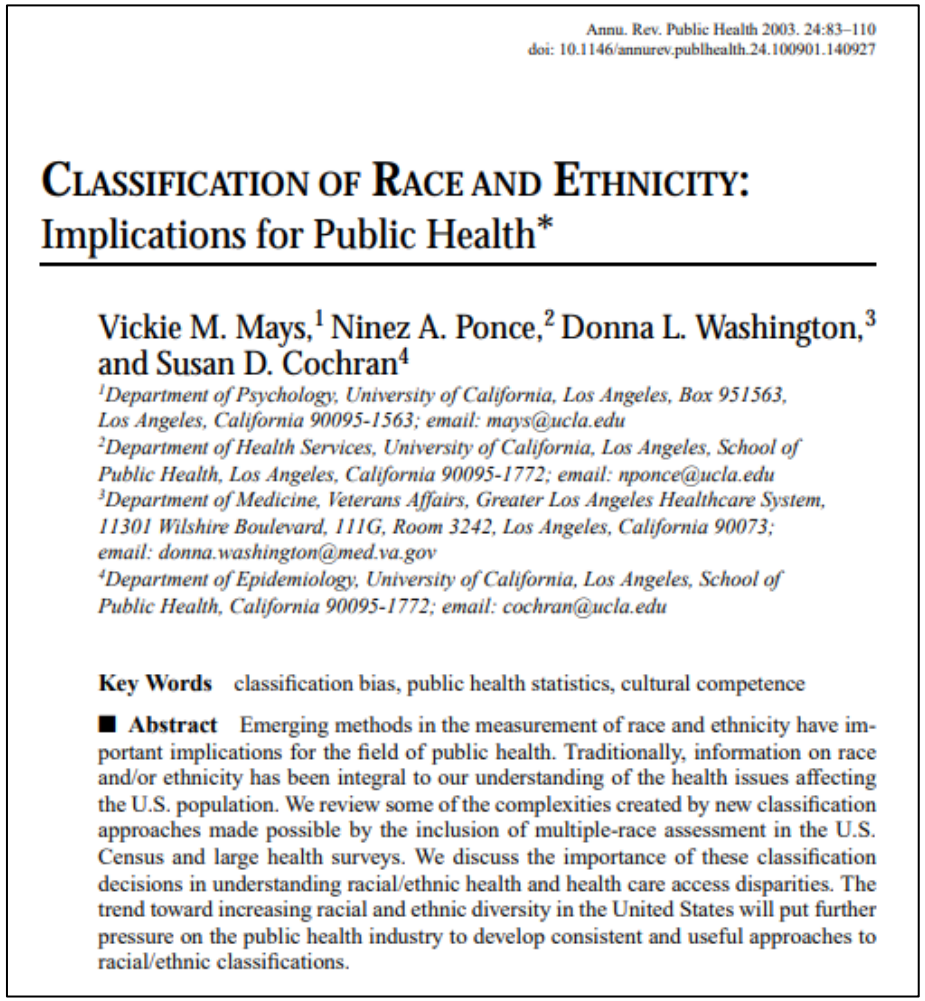
Grouping method

1. Persons with Hispanic ethnicity are grouped as Hispanic, regardless of race.
2. For the remaining records, persons with reported race and non-Hispanic, unknown, or missing ethnicity, are grouped as race category, non-Hispanic.
3. Persons with missing or unknown race and missing or unknown or non-Hispanic ethnicity are excluded.

- Excludes those who have selected multiple race categories
- Prioritize categorization of Hispanic ethnicity

https://www.cdc.gov/mmwr/volumes/70/wr/pdfs/mm7032a2-H.pdf

Standards to organize data for people who have selected multiple race and ethnicity categories are limited



Annu. Rev. Public Health 2003. 24:83–110
doi: 10.1146/annurev.publhealth.24.100901.140927

CLASSIFICATION OF RACE AND ETHNICITY: Implications for Public Health*

Vickie M. Mays,[1] Ninez A. Ponce,[2] Donna L. Washington,[3] and Susan D. Cochran[4]

[1]*Department of Psychology, University of California, Los Angeles, Box 951563, Los Angeles, California 90095-1563; email: mays@ucla.edu*
[2]*Department of Health Services, University of California, Los Angeles, School of Public Health, Los Angeles, California 90095-1772; email: nponce@ucla.edu*
[3]*Department of Medicine, Veterans Affairs, Greater Los Angeles Healthcare System, 11301 Wilshire Boulevard, 111G, Room 3242, Los Angeles, California 90073; email: donna.washington@med.va.gov*
[4]*Department of Epidemiology, University of California, Los Angeles, School of Public Health, California 90095-1772; email: cochran@ucla.edu*

Key Words classification bias, public health statistics, cultural competence

■ **Abstract** Emerging methods in the measurement of race and ethnicity have important implications for the field of public health. Traditionally, information on race and/or ethnicity has been integral to our understanding of the health issues affecting the U.S. population. We review some of the complexities created by new classification approaches made possible by the inclusion of multiple-race assessment in the U.S. Census and large health surveys. We discuss the importance of these classification decisions in understanding racial/ethnic health and health care access disparities. The trend toward increasing racial and ethnic diversity in the United States will put further pressure on the public health industry to develop consistent and useful approaches to racial/ethnic classifications.

- Called for standards for organizing the multi-select data
- Classify the multi-select into the category with lowest numbers or “rarest” in population estimates

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3681827/

Moving forward

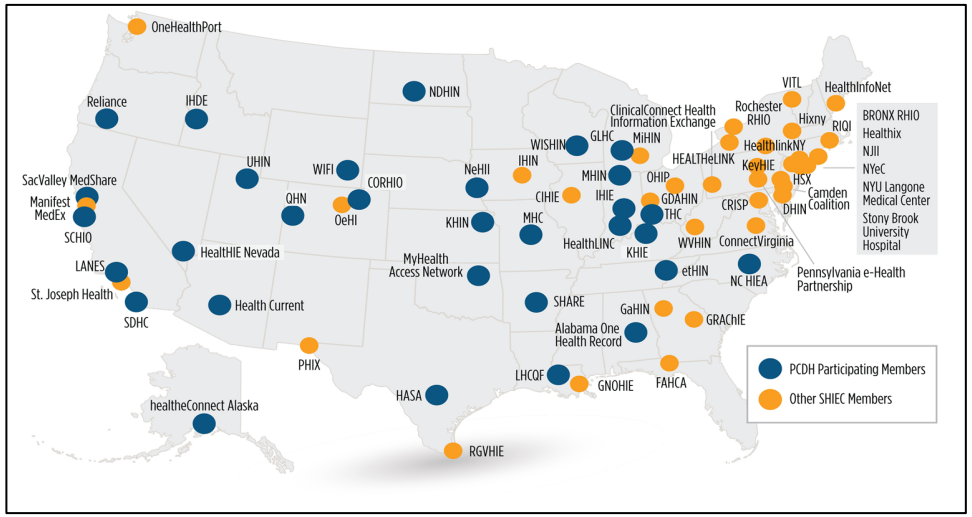
- Need standard for organizing data for people who select more than one category
  - Clinical care - data are shared across many different entities used for population health estimates
  - Research - sharing and aggregation across research studies, e.g., EHR-based studies
- Reconsider the use of these categories in clinical system
  - Patients social experience in care -- racism and discrimination in healthcare
  - Additional measures: self-reported socially assigned race and ethnicity in the EHR

What do we know about patient and community perspectives on data sharing?

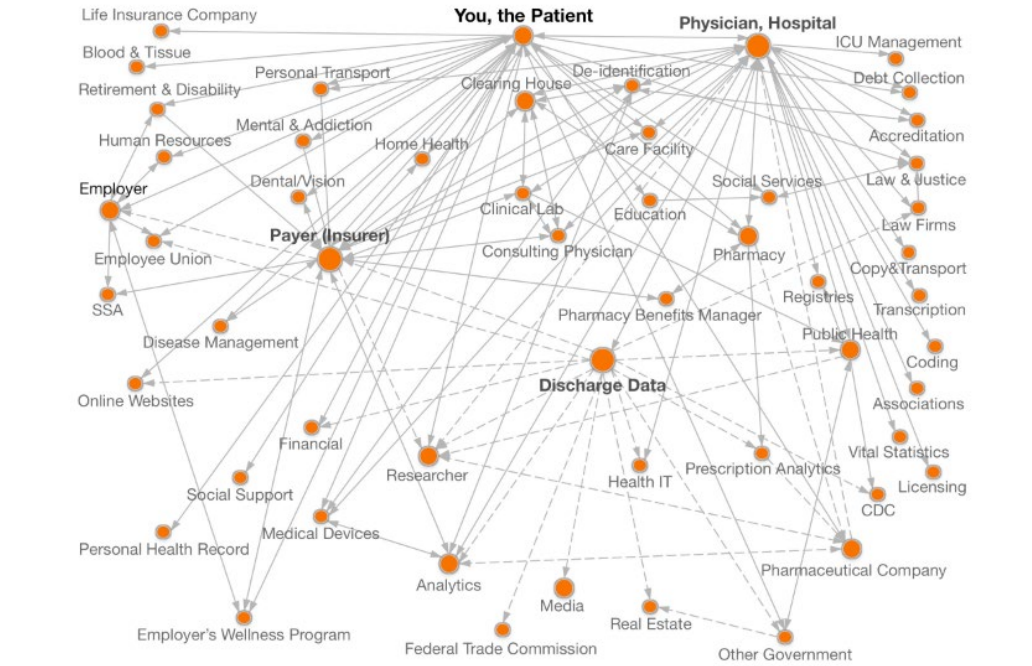
Landscape of collection and use has changed

Greater interoperability and ability to share patient records within and between healthcare entities across institutional and geopolitical borders

https://datavant.com/



https://www.healthit.gov/sites/default/files/page/2020-02/DavidKendrickSHIEC.pdf



https://thedatamap.org/

Patient perspectives on EHR collection and use

- Comfort with sharing their information for clinical care (Kim, K., 2015)
  - 56% were likely to agree to share medical information electronically with different places they receive medical care
  - Individual control, knowledge about who has access
- In current environment of data sharing, research on people's comfort in the collection and (re-) use of information in EHRs is nascent, related to people demographic information and social factors

Kim, Katherine K., Jill G. Joseph, and Lucila Ohno-Machado. "Comparison of consumers' views on electronic data sharing for healthcare and research." Journal of the American Medical Informatics Association 22.4 (2015): 821-830.

Moving forward

- Partner with minoritized communities on how we collect and use these data in the EHR for health equity
- Mechanism to audit and report how data are collected and used, who has access

Concluding thoughts

- Need a data justice framework and community-engagement to guide the use of race and ethnicity data in the EHR
  - Partner across disciplines and with patients and communities
  - Develop mechanism and measures of accountability to ensure collective benefit
    - Track use and sharing of race and ethnicity data
    - Share data back with communities
    - Use these data to intervene on effects of racism in healthcare and health

K. Wang, T Cullen, M Nunez-Smith. Centering Equity In The Design And Use Of Health Information Systems: Partnering With Communities On Race, Ethnicity, And Language Data, " Health Affairs, May 19, 2021.DOI: 10.1377/hblog20210514.126700

Acknowledgements

Noah Brazer

Sakinah Carter Suttiratana, PhD MPH MBA

Deron Galusha, MS

Corrine Liu, MPH

Marcella Nunez Smith, MD MHS

Carol Oladele, PhD MPH

Tara Rizzo, MPH

Baylah Tessier Sherman, MPH

DataHaven

CT Health Foundation, CT REL network members, Global Health Leadership Initiative

Yale Cancer Center/Smilow Cancer Hospital, Community Advisory Board