POLICY FORUM: PEER-REVIEWED ARTICLE

# What Are High-Quality Race and Ethnicity Data and How Are They Used in Health Equity Research?

Christopher T. Rentsch, PhD, FISPE Moneeza K. Siddiqui, PhD, MPH and Rohini Mathur, PhD, MS

## Abstract

The COVID-19 pandemic changed public awareness of the importance of high-quality race and ethnicity data for identifying and redressing widely documented racial and ethnic health inequity. This article emphasizes the importance of high-quality race and ethnicity data in health equity research, as highlighted by the COVID-19 pandemic. The article defines what constitutes high-quality race and ethnicity data, discusses the challenges in using these data, and provides 2 case studies that illustrate their role in identifying and redressing health inequity. Finally, this article advocates for the use of accurate, standardized, and granular data and highlights the need for community engagement and trust building to improve data quality and research outcomes.

## What Are Race and Ethnicity Data?

Race and ethnicity classifications reflect how particular groups of people have been racialized—that is, how their racial or ethnic identity has been shaped by historical and political forces. In particular, the way racial and ethnic groups are defined depend on social, cultural, political, and geographical context. Although the terms *race* and *ethnicity* have evolved over time, race has historically referred to broad categories of people that are divided arbitrarily based on ancestral origin and physical characteristics.[1] The United States (US) Census Bureau acknowledges that race is "a social definition … and not an attempt to define race biologically, anthropologically, or genetically."[2] In the US, ethnicity has historically referred to a person's cultural identity (e.g., language, customs, religion)—namely, as Hispanic or Latino, Latina, or Latinx.[3] In the United Kingdom (UK), however, the term *ethnicity* encompasses both of the above-mentioned concepts and is defined as the "various ways in which a person may choose to define their ethnic group… include[ing] common ancestry, elements of culture, identity, religion, language and physical appearance."[4] While the concepts of race and ethnicity are broad social constructs, they do not preclude the existence of biological or genetic variation that may affect health outcomes.[5] In this article, we use both terms—*race* and *ethnicity*—to refer to these social constructs, in line with recent proposals to use unified race and ethnicity categories.[6]

The COVID-19 pandemic changed public awareness of the importance of high-quality race and ethnicity data for identifying and redressing widely documented racial and ethnic health inequity. In health equity research, concepts of race and ethnicity can be thought of as proxies for structural and individual racism and discrimination.[7] In turn, research findings on racial or ethnic health differences, typically reported at a group or community level, are often a proxy for a range of health determinants, including—but not limited to—education, income, employment, housing, beliefs and behaviors, language and culture, and embodied experiences of racism and discrimination.[1,8] Thus, high-quality data on race and ethnicity can be a key first step to quantify health inequalities that ultimately are used as a basis for policy aimed at redressing them. In this article, we define what constitutes high-quality race and ethnicity data, discuss the challenges in using these data, and provide 2 case studies that illustrate their role in identifying and redressing health inequity.

## Characteristics of High-Quality Race and Ethnicity Data

Accurate and comprehensive data on race and ethnicity are critical for conducting effective health equity research to guide policy development. Essential characteristics of high-quality race and ethnicity data include high levels of completeness, self-reported collection, consistency, and granularity, as described below.

As with any data captured in routine health care settings, the completeness of data is related to health care usage and access, even in countries where health care is free at the point of access. Despite universal primary health care, certain population groups, such as migrants, attend primary care less frequently.[9] These important differences in access can greatly affect the completeness of race and ethnicity data, limiting our ability to redress disparities in populations often with the greatest health care need. The self-report of an individual's own racial or ethnic identity (as opposed to data recorded by an observer based on visual assessment or other indirect methods) is essential for accuracy.[10,11] While an individual's identity may not necessarily fit with the standard categories available to choose from, the use of consistent and standardized categories during collection and in published research minimizes discrepancies, enhances comparability, and allows for monitoring patterns over time. Greater granularity in racial and ethnic categories allows for better representation of racial and ethnic identities, provided analyses avoid combining relatively smaller groups into an "other" category that potentially obscures inequity. The quantity and validity of standard ethnic categories may evolve over time to reflect the changing ethnic makeup of a population. For example, the "mixed" ethnicity group is the largest growing ethnic group in the UK[12] and the US[13], and more granular breakdowns of this high-level, catch-all group will be essential for identifying the needs of the population over the long-term.

## Pandemic-Prompted Change

The COVID-19 pandemic has highlighted and exacerbated racial and ethnic inequity in health care and health outcomes.[14] Our understanding of these inequity was made possible by research leveraging routinely collected race and ethnicity data available in health care records and insurance claims databases. While several countries[15–17] recognize the importance of collecting race and ethnicity data, others consider the collection of such data illegal, making it impossible to directly quantify and redress inequity in these settings.[18,19] Collection of race and ethnicity data is an imperfect system, and current practices often suffer from inconsistencies in self-reported collection, standardization, and granularity of categories. However, these shortcomings should not preclude the use of existing race and ethnicity data to examine patterns in the health needs of minoritized populations.

The pandemic was a catalyst for change in research culture. The urgent need for responsive research led to widespread changes in how we use, share, and think about data. First, the pandemic resulted in initiatives (as demonstrated in the cases below) that improved the speed, safety, and transparency of research. Second, it placed inequalities research in a global spotlight. Early in the pandemic, press reports suggested that racially and ethnically minoritized groups were disproportionately affected by COVID-19 relative to their White counterparts.[20–22] Hypotheses included excess occupational exposure to the SARS-CoV-2 virus, greater barriers in accessing health care, and lack of culturally and linguistically appropriate public health communications.[23,24] There was a clear and urgent need to formally evaluate the potential for racial and ethnic inequity associated with the pandemic. Third, the pandemic led to novel collaborations across sectors and disciplines, including community partnerships and engagement. For example, Latino communities in California engaged in community-academic partnerships to develop culturally appropriate health interventions addressing testing barriers.[25] Fourth, it required researchers to facilitate public understanding to help narrow the “trust gap” between researchers and the public concerning how people’s health and administrative data are used for research.[26.] These changes in research practice hold promise for more rapidly translating scientific research into policy aimed at redressing health inequalities.

## Cases

Below, we provide 2 use cases that demonstrate the benefits and challenges of using race and ethnicity data to identify and redress inequity in health care utilization and outcomes. The cases we selected represent health care systems in the US and UK that offer care largely free of charge, thereby minimizing significant cost barriers to health care utilization. However, inequity in access to health care remain in both systems.[27,28] Disentangling the impact of health care access from observed inequity in health outcomes remains a challenge, as any underrepresentation of marginalized groups in the data can compromise the ability to accurately assess and redress health inequity.

*Case 1: racial and ethnic disparities in COVID-19 pandemic in the US and UK.* In the US, we highlight research leveraging longitudinal electronic health record data from the Department of Veterans Affairs (VA). The VA is the largest integrated health care system in the US and provides comprehensive health care to more than 9 million Veterans annually nationwide at over 1300 points-of-care.[29] Since 2003, the VA has routinely collected self-reported race and ethnicity data during intake and at outpatient and inpatient visits.[11] In the UK, we highlight research conducted using OpenSAFELY,[30] a novel software platform developed on behalf of NHS England to support rapid, responsive research on COVID-19. At its inception in 2020, OpenSAFELY included electronic health records that contained self-reported ethnicity[31] for 25 million people, covering 40% of the English population.[32]

With VA data, researchers identified stark disparities in testing positive[33–35] and COVID-19 hospitalizations[36] among racial and ethnic minoritized groups. However, among those who tested positive, there were no observed disparities in subsequent mortality,[33] which has been attributed to the care received in the VA health care system, as health disparities in the VA tend to be smaller than in the private sector.[37] Nevertheless, at a population level the substantial excess burden of SARS-CoV-2 infection among racially and ethnically minoritized groups inevitably translated to excess mortality in these communities in the US[38] and UK.[39] In the US, American Indian and Alaska Native (AI/AN) patients "experienced the largest absolute and relative increases in mortality during the pandemic," although they represented only 1% of the VA population.[40] The OpenSAFELY studies found similar ethnic disparities in testing positive, hospitalization, and mortality.[41] In the UK data, these data were used to additionally identify factors such as living in deprived areas[42] and residing in large, multigenerational households[43] associated with SARS-CoV-2 infection and mortality. Thanks to large sample sizes, researchers were able to undertake comparisons among more granular ethnicity groups, which identified widening inequalities in COVID-19 mortality among South Asian groups, especially the Bangladeshi community, in the second wave of the pandemic. These findings led to further work in which the crude household size variable was redefined as a measure of multigenerational living. This work showed that 66% of people of South Asian ethnicity live in multigenerational households compared to 23% of White groups and 49% of Black groups and that multigenerational living and living alone were both associated with increased risk of COVID-19.[43] In both countries, however, the lack of data on wider social determinants of health, such as employment and contact patterns, in large-scale electronic health record systems limited investigating these factors further.

Despite these limitations, the rapid, responsive way of working during the pandemic meant that researchers in both countries were collaborating in large, multidisciplinary teams, enabling rapid transformation of research findings into responsive policy recommendations, including for tailored, culturally responsive public health messaging concerning prevention and, eventually, vaccination. For example, the VA created a COVID-19 Equity Dashboard to track and visualize infection and vaccination rates by race and ethnicity and other demographic factors, enabling targeted outreach and intervention.[44] Additionally, the VA conducted virtual listening sessions between Veterans of color and demographic-matched professionals to increase vaccination rates and address vaccine hesitancy by building trust and explicitly redressing historical injustices.[44] In the UK, targeted communication and engagement strategies, such as leveraging local influencers through the Community Champions scheme and utilizing flexible deployment models that support vaccinations during religious events and in places of worship, were essential to improving vaccine uptake among ethnic minorities and combatting misinformation.[45]

To maximize transparency and trust in its research, each study conducted using the OpenSAFELY platform is required to preregister a complete study protocol and publicly share all code that extracts and analyzes data.[46,47] This transparency aims to assure all stakeholders, including patients, professionals, and policy makers, that data were used as intended and handled and interpreted appropriately.

*Case 2: using ethnicity data to develop targeted public health interventions.* For over 30 years, the Clinical Effectiveness Group (CEG) at Queen Mary, University of London, has utilized electronic health record data to generate valuable insights and innovations, thereby facilitating health and social care improvements. The CEG enhances learning health systems in one of London's most diverse and deprived areas, the borough of Tower Hamlets. By employing a cycle of analysis, feedback, and interaction, the CEG effectively bridges research, policy, and practice, driving public health advancements and reducing inequalities.

One key example of the learning health system at work is that of redressing ethnic inequities in measles mumps and rubella (MMR) vaccination. It was found that "between 2006 and 2008 … Tower Hamlets had the highest rates of confirmed measles [in the UK], with 24 cases per 100 000 … compared with a national figure of 2 per 100 000."[48] Using routinely collected primary care data, the CEG as able to demonstrate significant ethnic inequities in MMR uptake. In Tower Hamlets, focus group work with Somali parents suggested that MMR vaccine uptake was low partly on account of safety concerns related to autism. Thanks to high-quality ethnicity recording (which was over 97% complete for children under 5), the researchers were able to analyze data for the Somali group separately from the broader ethnic category of Black African/Caribbean.

By 2011, Tower Hamlets had virtually achieved herd immunity and had the highest rates of MMR vaccination in London, thanks to efforts that were responsive to the local context.[48] The CEG demonstrated that achieving herd immunity for childhood vaccinations was an achievable goal in an ethnically and socially diverse population. The high-quality ethnicity data available to researchers allowed them “to identify characteristics of the difficult to reach groups, including significant differences in uptake across different ethnicities.”[48]

Changes in management and the withdrawal of financial incentives meant that the gains were not sustained long-term. Ten years later, MMR immunization rates in London dropped to levels disproportionately lower than the rest of the UK, partly due to the pandemic.[49] Inequity widened, prompting renewed efforts to reach herd immunity for MMR. In February 2022, the CEG launched a quality improvement program to redress falling rates of childhood immunizations. Research is now underway to fully evaluate the program, which will generate the evidence base to inform practice and policy going forward.[50] One suggested policy action is to include national measures to tackle these inequities by financially incentivizing general practitioners to deliver timely routine childhood vaccinations in primary care.[49–51]

## Current Key Challenges

Achieving representative data collection presents significant challenges, especially in diverse populations in which socioeconomic inequity, access to health care, and geographic location can influence data quality and availability. It is further complicated in systems where race and ethnicity data collection can be skewed by the nature of healthcare provision. Although health care systems like the VA or the UK’s National Health Service are largely free at the point of contact, those who are marginalized might be less likely to interact with health care systems and be represented in the data.

While the above cases constitute positive examples of using existing large-scale race and ethnicity data, data injustices remain. For example, the term *data genocide* has been used to describe the lack of AI/AN data available in the US during the pandemic.[52] As a result, AI/AN communities exercised communal ownership of health data to drive public health responses tailored to their specific needs.[53] Greater community engagement is crucial in redressing health inequity and building trust between researchers and marginalized communities.

To overcome these challenges, as a start, we point to recent guidance on the reporting of race and ethnicity in scientific research.[54] We also note a call for action to bring about data justice "regarding the reporting and analysis of publicly-funded work involving racialized groups."[1]

## Conclusion

Ensuring high-quality race and ethnicity data through self-reported collection of standardized, granular categories is crucial for meaningful analysis aimed at identifying health inequity. Given the underlying factors inherent in racial and ethnic group classifications, analyzing inequity can yield crucial insights into health patterns and serve as a critical basis for redressing health inequity.

## References

1. Krieger N. Structural Racism, Health Inequities, and the Two-Edged Sword of Data: Structural Problems Require Structural Solutions. *Front Public Health*. 2021;9:655447. doi:10.3389/fpubh.2021.655447

2. United States Census Bureau. About the Topic of Race. Accessed June 27, 2024. https://www.census.gov/topics/population/race/about.html

3. Stamper K. Conscious Style Guide. Why we confuse race and ethnicity: a lexicographer's perspective. February 13, 2019. Accessed May 8, 2021. https://consciousstyleguide.com/why-we-confuse-race-ethnicity-lexicographers-perspective/

4. Office for National Statistics. Ethnic group, England and Wales: Census 2021. Accessed June 27, 2024. https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/bulletins/ethnicgroupenglandandwales/census2021

5. Borrell LN, Elhawary JR, Fuentes-Afflick E, et al. Race and Genetic Ancestry in Medicine — A Time for Reckoning with Racism. Malina D, ed. *N Engl J Med*. 2021;384(5):474-480. doi:10.1056/NEJMms2029562

6. Flores G. Language Barriers and Hospitalized Children: Are We Overlooking the Most Important Risk Factor for Adverse Events? *JAMA Pediatr*. 2020;174(12):e203238. doi:10.1001/jamapediatrics.2020.3238

7. Lett E, Asabor E, Beltrán S, Cannon AM, Arah OA. Conceptualizing, Contextualizing, and Operationalizing Race in Quantitative Health Sciences Research. *Ann Fam Med*. 2022;20(2):157-163. doi:10.1370/afm.2792

8. World Health Organization. Social determinants of health. Accessed June 27, 2024. https://www.who.int/health-topics/social-determinants-of-health#tab=tab_1

9. Zhang CX, Boukari Y, Pathak N, et al. Migrants' primary care utilisation before and during the COVID-19 pandemic in England: An interrupted time series analysis. *Lancet Reg Health - Eur*. 2022;20:100455. doi:10.1016/j.lanepe.2022.100455

10. Bastos JL, Peres MA, Peres KG, Dumith SC, Gigante DP. Socioeconomic differences between self- and interviewer-classification of color/race. *Rev Saude Publica*. 2008;42(2):324-334. doi:10.1590/s0034-89102008000200019

11. Sohn MW, Zhang H, Arnold N, et al. Transition to the new race/ethnicity data collection standards in the Department of Veterans Affairs. *Popul Health Metr*. 2006;4(1):7. doi:10.1186/1478-7954-4-7

12. Migration Museum. Census reveals new chapter in story of mixed-race Britain. December 7, 2022. Accessed August 15, 2024. https://www.migrationmuseum.org/census-reveals-new-chapter-in-story-of-mixed-race-britain/#:~:text=Source%3A%20Office%20for%20National%20Statistics%20–%20Census%202021&text=The%20number%20of%20people%20identifying%20as%20'White%20and%20Asian'%20rose,%3A%20up%2061%25%20to%20467%2C113.

13. Pew Research Center. Multiracial in America. June 11, 2015. Accessed August 15, 2024. https://www.pewresearch.org/social-trends/2015/06/11/multiracial-in-america/

14. Katikireddi SV, Lal S, Carrol ED, et al. Unequal impact of the COVID-19 crisis on minority ethnic groups: a framework for understanding and addressing inequalities. *J Epidemiol Community Health*. 2021;75(10):970-974. doi:10.1136/jech-2020-216061

15. Ministry of Social Development. Improving how we report ethnicity. Accessed June 27, 2024. https://www.msd.govt.nz/about-msd-and-our-work/tools/how-we-report-ethnicity.html

16. Deb S, Sud M, Coburn N, et al. Race and Ethnicity Research in Cardiovascular Disease in Canada: Challenges and Opportunities. *Can J Cardiol*. 2024;40(6):1172-1175. doi:10.1016/j.cjca.2024.02.001

17. Anjana RM, Unnikrishnan R, Deepa M, et al. Metabolic non-communicable disease health report of India: the ICMR-INDIAB national cross-sectional study (ICMR-INDIAB-17). *Lancet Diabetes Endocrinol*. 2023;11(7):474-489. doi:10.1016/S2213-8587(23)00119-5

18. Migration Policy Group. Combating racial and ethnic discrimination: taking the European legislative agenda further. Accessed June 27, 2024. https://www.migpolgroup.com/_old/wp-content/uploads/2016/10/81.CombatingRacialandEthnicDiscrimination-TakingtheEuropeanLegislativeAgendaFurther_03.02.pdf

19. Al-Zubaidi Y. Racial and Ethnic Statistics in Sweden: Has the Socialization Process Started Yet? Accessed June 27, 2024. https://scandinavianlaw.se/pdf/68-18.pdf

20. Aldridge RW, Lewer D, Katikireddi SV, et al. Black, Asian and Minority Ethnic groups in England are at increased risk of death from COVID-19: indirect standardisation of NHS mortality data. *Wellcome Open Res*. 2020;5:88. doi:10.12688/wellcomeopenres.15922.2

21. Reyes C, Husain N, Gutowski C, St. Clair S, Pratt G. Chicago's coronavirus disparity: black Chicagoans are dying at nearly six times the rate of white residents, data show. *Chicago Tribune*. April 7, 2020.

22. Thebault R, Tran A, Williams V. The coronavirus is infecting and killing black Americans at an alarmingly high rate. *Washington Post*. April 7, 2020.

23. Fothergill A, Maestas EG, Darlington JD. Race, ethnicity and disasters in the United States: a review of the literature. *Disasters*. 1999;23(2):156-173. doi:10.1111/1467-7717.00111

24. Webb Hooper M, Nápoles AM, Pérez-Stable EJ. COVID-19 and Racial/Ethnic Disparities. *JAMA*. 2020;323(24):2466-2467. doi:10.1001/jama.2020.8598

25. Garibay KK, Durazo A, Vizcaíno T, et al. Lessons from Two Latino Communities Working with Academic Partners to Increase Access to COVID-19 Testing. *Prog Community Health Partnersh Res Educ Action*. 2024;18(1):1-9.

26. Mathur R, Rentsch CT, Venkataraman K, et al. How do we collect good-quality data on race and ethnicity and address the trust gap? *Lancet Lond Engl*. 2022;400(10368):2028-2030. doi:10.1016/S0140-6736(22)02490-4

27. Ajayi Sotubo O. A perspective on health inequalities in BAME communities and how to improve access to primary care. *Future Healthc J*. 2021;8(1):36-39. doi:10.7861/fhj.2020-0217

28. Ward RE, Nguyen XMT, Li Y, et al. Racial and Ethnic Disparities in U.S. Veteran Health Characteristics. *Int J Environ Res Public Health*. 2021;18(5):2411. doi:10.3390/ijerph18052411

29. US Department of Veterans Affairs. Veterans Health Administration: About VHA. Accessed August 15, 2024. https://www.va.gov/health/aboutvha.asp

30. OpenSAFELY. Secure analytics platform for NHS electronic health records. Accessed June 27, 2024. https://www.opensafely.org/

31. Hull SA, Mathur R, Badrick E, Robson J, Boomla K. Recording ethnicity in primary care: assessing the methods and impact. *Br J Gen Pract*. 2011;61(586):e290-e294. doi:10.3399/bjgp11X572544

32. The OpenSAFELY Collaborative, Andrews CD, Mathur R, et al. Consistency, completeness and external validity of ethnicity recording in NHS primary care records: a cohort study in 25 million patients' records at source using OpenSAFELY. *BMC Med*. 2024;22(1):288. doi:10.1186/s12916-024-03499-5

33. Rentsch CT, Kidwai-Khan F, Tate JP, et al. Patterns of COVID-19 testing and mortality by race and ethnicity among United States veterans: A nationwide cohort study. *PLoS Med*. 2020;17(9):e1003379. doi:10.1371/journal.pmed.1003379

34. Ferguson JM, Abdel Magid HS, Purnell AL, Kiang MV, Osborne TF. Differences in COVID-19 Testing and Test Positivity Among Veterans, United States, 2020. *Public Health Rep Wash DC 1974*. 2021;136(4):483-492. doi:10.1177/00333549211009498

35. Ferguson JM, Justice AC, Osborne TF, Magid HSA, Purnell AL, Rentsch CT. Geographic and temporal variation in racial and ethnic disparities in SARS-CoV-2 positivity between February 2020 and August 2021 in the United States. *Sci Rep*. 2022;12(1):273. doi:10.1038/s41598-021-03967-5

36. Razjouyan J, Helmer DA, Li A, et al. Differences in COVID-19-Related Testing and Healthcare Utilization by Race and Ethnicity in the Veterans Health Administration. *J Racial Ethn Health Disparities*. 2022;9(2):519-526. doi:10.1007/s40615-021-00982-0

37. Peterson K, Anderson J, Boundy E, Ferguson L, McCleery E, Waldrip K. Mortality Disparities in Racial/Ethnic Minority Groups in the Veterans Health Administration: An Evidence Review and Map. *Am J Public Health*. 2018;108(3):e1-e11. doi:10.2105/AJPH.2017.304246

38. Weinberger DM, Rose L, Rentsch C, et al. Excess Mortality Among Patients in the Veterans Affairs Health System Compared With the Overall US Population During the First Year of the COVID-19 Pandemic. *JAMA Netw Open*. 2023;6(5):e2312140. doi:10.1001/jamanetworkopen.2023.12140

39. Strongman H, Carreira H, De Stavola BL, Bhaskaran K, Leon DA. Factors associated with excess all-cause mortality in the first wave of the COVID-19 pandemic in the UK: A time series analysis using the Clinical Practice Research Datalink. *PLoS Med*. 2022;19(1):e1003870. doi:10.1371/journal.pmed.1003870

40. Weinberger DM, Bhaskaran K, Korves C, et al. Excess mortality in US Veterans during the COVID-19 pandemic: an individual-level cohort study. *Int J Epidemiol*. 2023;52(6):1725-1734. doi:10.1093/ije/dyad136

41. Mathur R, Rentsch CT, Morton CE, et al. Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the OpenSAFELY platform. *Lancet Lond Engl*. 2021;397(10286):1711-1724. doi:10.1016/S0140-6736(21)00634-6

42. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature*. 2020;584(7821):430-436. doi:10.1038/s41586-020-2521-4

43. Wing K, Grint DJ, Mathur R, et al. Association between household composition and severe COVID-19 outcomes in older people by ethnicity: an observational cohort study using the OpenSAFELY platform. *Int J Epidemiol*. 2022;51(6):1745-1760. doi:10.1093/ije/dyac158

44. Leder SC, List JM, Chandra R, Jones KT, Moy E. VA Research and Operations Uniting to Combat COVID-19 Inequities. *Health Equity*. 2023;7(1):296-302. doi:10.1089/heq.2023.0002

45. UK Government. Third quarterly report on progress to address COVID-19 health inequalities. September 3, 2021. Accessed June 27, 2024. https://www.gov.uk/government/publications/third-quarterly-report-on-progress-to-address-covid-19-health-inequalities/third-quarterly-report-on-progress-to-address-covid-19-health-inequalities

46. OpenSAFELY GitHub. Accessed February 27, 2024. https://github.com/OpenSAFELY

47. OpenSAFELY Jobs Server. Accessed February 27, 2024. https://jobs.opensafely.org/

48. Cockman P, Dawson L, Mathur R, Hull S. Improving MMR vaccination rates: herd immunity is a realistic goal. *BMJ*. 2011;343:d5703. doi:10.1136/bmj.d5703

49. Firman N, Marszalek M, Gutierrez A, et al. Impact of the COVID-19 pandemic on timeliness and equity of measles, mumps and rubella vaccinations in North East London: a longitudinal study using electronic health records. *BMJ Open*. 2022;12(12):e066288. doi:10.1136/bmjopen-2022-066288

50. Marszalek M, Hawking MKD, Gutierrez A, et al. Implementation of a quality improvement programme using the Active Patient Link call and recall system to improve timeliness and equity of childhood vaccinations: protocol for a mixed-methods evaluation. *BMJ Open*. 2023;13(1):e064364. doi:10.1136/bmjopen-2022-064364

51. British Medical Association, NHS England. Update to the GP contract agreement 2020/21 - 2023/24. February 6, 2020. Accessed June 27, 2024. https://www.england.nhs.uk/wp-content/uploads/2020/03/update-to-the-gp-contract-agreement-v2-updated.pdf

52. Urban Indian Health Institute. Data Genocide of American Indians and Alaska Natives in COVID-19 Data. Accessed June 27, 2024. https://www.uihi.org/projects/data-genocide-of-american-indians-and-alaska-natives-in-covid-19-data/

53. Huyser KR, Horse AJY, Kuhlemeier AA, Huyser MR. COVID-19 Pandemic and Indigenous Representation in Public Health Data. *Am J Public Health*. 2021;111(S3):S208-S214. doi:10.2105/AJPH.2021.306415

54. Flanagin A, Frey T, Christiansen SL, AMA Manual of Style Committee. Updated Guidance on the Reporting of Race and Ethnicity in Medical and Science Journals. *JAMA*. 2021;326(7):621. doi:10.1001/jama.2021.13304

## AUTHOR INFORMATION

**Christopher T. Rentsch, PhD, FISPE** is an associate professor at the London School of Hygiene & Tropical Medicine (LSHTM) in England and an assistant professor (adjunct) at the Yale School of Medicine in New Haven, Connecticut. He obtained an MPH from Emory University and a PhD from LSHTM. Dr Rentsch specializes in the use of electronic health records to generate real-world evidence for the safety and effectiveness of medications, with a focus on quantifying disparities in medication receipt and outcomes.

**Moneeza K. Siddiqui, PhD, MPH** is a lecturer in genetic epidemiology with an interest in precision medicine and previously served as a principal investigator in precision medicine at the University of Dundee in Scotland, where she led research in pharmacogenetics and type 2 diabetes. She obtained an MPH from Columbia University in New York City and a PhD from the University of Dundee. Her research focuses on comparisons across ancestries, using genetics and multi-omics methods to understand the heterogenous presentation of type 2 diabetes in South Asians, particularly young-onset diabetes.

**Rohini Mathur, PhD, MS** is a professor and the chair of health data science at Queen Mary University of London in England, specializing in health inequalities research in the UK and globally. She is also the academic lead of the Clinical Effectiveness Group, a clinically driven and academically supported quality improvement and research center, where she leads research with local and international partners. She obtained an MS and PhD from the London School of Hygiene and Tropical Medicine. Her research aims to generate an evidence base to inform tailored approaches to the management of cardiometabolic disease by harnessing global data on pathophysiology, clinical outcomes, and treatment response in ethnically and geographically diverse populations.

### Conflict of Interest Disclosure

Dr Mathur reports is affiliated with the Genes & Health study, which is partly funded (including salary contributions) by a Life Sciences Consortium comprising Astra Zeneca, Bristol-Myers Squibb, GlaxoSmithKline Research and Development Limited, Maze Therapeutics, Merck Sharp & Dohme, Novo Nordisk, Pfizer, and Takeda Development Center Americas. The other authors disclosed no conflicts of interest.