USING RACE & RELATED DATA

In the context of your ICES research question

Race affects health because it influences access to resources and experiences of everyday life, all of which have physiological, behavioural, and other impacts. As we describe below, there are many theoretical dimensions of race, including whether it is self-reported (one's own assessment of one's race) or observed or socially assigned (one's race as perceived by others).[1,2] There are also related constructs, such as ethnicity. On the other hand, by and large, the data held at ICES (and most administrative data) only contain information on some of these constructs. The main issue is to **understand what you want to measure and why,** and then to see **if and how that fits into the reality of the data you have**.[3] Although there are many advantages to reporting on race and related data, the harms of misusing these data or miscategorizing people in the process can also be significant.

For more information, explore the inventory of race and related data within the ICES repository (listed under the Tools section of our **Resources for Anti-Racist Research**)

*This document does not address Indigenous identifiers. ICES has distinct approaches and policies for Indigenous data governance in partnership with Indigenous leadership organizations.
*Please contact **Indigenous Partnerships, Data and Analytics** for guidance*

FOUNDATIONAL TERMS

Race is a way that society groups people into categories based on phenotypic characteristics, such as skin colour and facial features. It can vary by who is assessing race (i.e., self-reported versus observed/socially-assigned).

Self-reported race and ethnicity offer value as a gold standard for affirming individual identity, and centering empowerment and self-determination.[4] Due to most administrative and survey data containing self-reported data, the scientific literature commonly uses self-reported race as a proxy for observed race (e.g., in analyses of experiences of racism or discrimination). For most people, their self-reported race matches observed race, but it is important to note that, in some cases, self-reported race and observed race may be discordant, which can introduce error. Both societal perception and individual self-perception of racial categories, their names, and who is in or out of a given category can change across time and is different across countries. Race is often mistakenly thought to indicate genetic groupings, but population genetics research shows this is false: race cannot even be deduced by looking at someone's DNA, and there is often more genetic variation within racial groups than between them. It is also not (just or even mainly) a cultural category.

Ethnicity is also a way that society groups people into categories, largely based on family geographic origin, language, religion, nationality, and other characteristics. Two people of the same racial group could have different ethnicities, or two people with the same ethnicity may have different racial groups.

Ethnicity is also often thought of as a genetic grouping, which it is not. It is also often thought of as a cultural category, which it might be. It also might be a category that shapes power, resources, and everyday experiences. Defining and measuring ethnicity is complex and may intersect with related constructs such as religious identity. For example, Jewish ethnic or cultural origins are not the same as Jewish religion. If "Jewish" is absent from available response categories for ethnicity or ancestry, Jews may choose not to respond or may choose a category that doesn't accurately or completely reflect their self-perception.[5] Moreover, various contextual factors may shape how people respond to this question, in addition to what they feel is their 'true' identification.

IC/ES

Though race and ethnicity have different technical definitions, our society tends to categorize people in ways that do not map into these terms in a mutually exclusive way, so some suggest unifying the constructs into an aggregate term ("race and ethnicity").[6]

For instance, "South Asian" is a grouping that might be considered both a "race" (phenotypically defined) and an "ethnicity" (culturally defined).

Ancestry must be distinguished from "race" and "ethnicity" (categories that are socially constructed and do not represent genetically or biologically distinct homogenous groups).[3]

If a genetic condition is associated with a specific ancestral region (e.g., prevalence of sickle cell trait in regions whose conditions support malaria-carrying *Anopheles* mosquitoes), this should be articulated explicitly to avoid conflation with socially-constructed race categories that are poorer proxies for region.

SOURCES OF RACE & RELATED DATA AT ICES

The province of Ontario does not routinely collect data on race or related constructs; however, data on race and related constructs are available in some data holdings at ICES, with most (but not all) of these data being self-reported. It's important to note that ICES' race-related variables may not necessarily map neatly into mutually exclusive source categories and some data holdings may collect race-related variables using multiple methods, such as a combination of observer reported and self-reported. All data have limitations, and it's important to be transparent about limitations, including whether method of collection influences how the results are interpreted.[7]

THREE COMMON SOURCES OF RACE & RELATED DATA AT ICES

1. **The Canadian Community Health Survey (CCHS)** is a national health survey for which Ontario responses are linked to ICES data holdings. CCHS contains a variety of sociodemographic indicators, including a question on race/ethnicity that closely resembles the question on race/ethnicity asked during the Statistics Canada Census. The categorical choice set is fixed. Respondents can choose as many categories as apply to them, and there is also an 'other' category. The categorization is not exhaustive and there are categories that are excluded, particularly those that are interactions between race, ethnicity, and religious identity.

2. The Ontario Marginalization Index (ON-Marg) provides census-based, area-level information and has a stand-alone dimension of "racialized and newcomer populations." This index is based on both the percent of "racialized" (non-White) residents and the percent of immigrants in a census area (e.g., Dissemination Area, Census Tract, etc.). This index should **not** be used as a proxy for individual-level race/ethnicity. Rather, the index may help to identify, in an ecological analysis, whether health or health services vary across areas with different proportions of racialized and newcomer populations.

3. The Immigration, Refugees and Citizenship Canada (IRCC/CIC) Ontario Data provide information related to people's immigration characteristics, including one's country of birth (self-reported) and country of emigration. These variables should **not** be used as a proxy for race or ethnicity, as many racialized people in Ontario will **not** be accurately classified in this conflation. Currently, the IRCC "immigrant" data available at ICES are only for immigrants and refugees who landed in Ontario from 1985 onwards and do not include immigrants who landed earlier or who first landed in another province (and later moved to Ontario). Although ICES plans to eventually integrate data on immigrants who landed in other provinces, there will remain no data from years prior to 1985. As such, consider this category is more accurately labelled "Canadian-born/long-term resident" (instead of "nonimmigrant," which is not recommended). When appropriate, subclassifying "immigrants" is also encouraged to better contextualize the data: for example, by incorporating information on immigration category (e.g., refugee, economic immigrant), sociodemographics (e.g., country of birth/origin, neighbourhood income), and other information (e.g., recency of landing, language at time of immigration) and labeling the constructs as they are defined and not as proxies.

BOTTOM LINE: THINK CRITICALLY ABOUT WHAT YOU ARE TRYING TO MEASURE

Thinking critically about the distinct limitations of each data holding will promote the most appropriate use of the data we have. Using multiple dimensions of race also may aid in more nuanced examination of the effects of racism.[2] Developing a conceptual framework may aid in considering and communicating what race-related variables are serving as proxies.[7] Other recommended practices to reduce misinterpretation include using the most specific description of the source data possible (e.g., "socially-assigned race" vs "self-reported racial identity").

IC/ES

Select Journal Guidance on Reporting Race and Ethnicity Data in Research

3) Stanbrook MB, Salami B. CMAJ's new guidance on the reporting of race and ethnicity in research articles. *CMAJ.* 2023 February 13;195:E236-8. https://doi.org/10.1503/cmaj.230144

6) 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):621–627. https://doi.org/10.1001/jama.2021.13304

7) Duncan AF, Montoya-Williams D. Recommendations for Reporting Research About Racial Disparities in Medical and Scientific Journals. *JAMA Pediatr.* 2024;178(3):221–224. https://doi.org/10.1001/jamapediatrics.2023.5718

Chew M, Samuel D, the Lancet Group for Racial Equity (GRacE), Mullan Z, Kleinert S. The Lancet Group's new guidance to authors on reporting race and ethnicity. *The Lancet*. 2024;403(10442):2360–2361. https://doi.org/10.1016/s0140-6736(24)01081-x

Other Supporting Literature

1) Roth WD. The multiple dimensions of race. *Ethnic and Racial Studies*. 2016. 39(8), 1310–1338. https://doi.org/10.1080/01419870.2016.1140793

2) Sansone G, Adnan A, Fallon B, et al. Social Identification of Race in Canada. SocialWork.utoronto.ca. Published January 2023. Accessed March 29, 2023. https://socialwork.utoronto.ca/fraser-mustard-institute-for-human-development-fmihd-policy-bench/

4) Engagement, Governance, Access, and Protection (EGAP): A Data Governance Framework for Health Data Collected from Black Communities. Black Health Equity Working Group. Published 2021. Accessed 2022. https://blackhealthequity.ca/wp-content/uploads/2021/03/Report_EGAP_framework.pdf

5) Smith T, McLeish S. Technical report on changes in response related to the census ethnic origin question: Focus on Jewish origins, 2016 Census integrated with 2011 National Household Survey. *Statistics Canada: Ethnicity, Language and Immigration Thematic Series*. Published July 26, 2019. Accessed October 2024. https://www150.statcan.gc.ca/n1/pub/89-657-x/89-657-x2019009-eng.htm

Chou V. How Science and Genetics are Reshaping the Race Debate of the 21st Century. *Science in the News, Harvard University.* Published April 17, 2017. Accessed May 27, 2022. https://sitn.hms.harvard.edu/flash/2017/science-genetics-reshaping-race-debate-21st-century/

Darity WA, Dietrich J, Hamilton D. Bleach in the rainbow: Latin ethnicity and preference for whiteness. *Transforming Anthropology.* 2005;13(2):103-109. https://doi.org/10.1525/tran.2005.13.2.103

Jones CP, Truman BI, Elam-Evans LD, et al. Using "socially assigned race" to probe White advantages in health status. Ethn Dis. 2008;18(4):496–504. https://pubmed.ncbi.nlm.nih.gov/19157256/

Roth WD. Racial mismatch: the divergence between form and function in data for monitoring racial discrimination of Hispanics. Soc Sci Q. 2010;91(5):1288–311. https://doi.org/10.1111/j.1540-6237.2010.00732.x

White K, Lawrence JA, Tchangalova N, et al. Socially-assigned race and health: a scoping review with global implications for population health equity. Int J Equity Health. 2020;19. https://doi.org/10.1186/s12939-020-1137-5