



Pan-Canadian
Genome Library

Bibliothèque génomique
pancanadienne

PCGL Guidance Document for Participant Sociodemographic Data

Version 1.0. | May 02, 2025

Corresponds to *PCGL Data Dictionary* version 1.0_2025-05-05

Corresponds to *PCGL Reference Survey for Participant Sociodemographic Data* version 1.0



Acknowledgments

Production of this document is made possible by financial contributions from the Canadian Institutes of Health Research and Genome Canada. The views expressed herein do not necessarily represent the views of the Canadian Institutes of Health Research and Genome Canada.

How to cite this document:

Chan A, Hicks AI, Pelletier J, Pan-Canadian Genome Library, Bourque G and Bhérer C. *PCGL Guidance Document for Participant Sociodemographic Data Version 1.0*. Montréal, QC: PCGL; 2025.



REVISION HISTORY.....	4
INTRODUCTION	5
BACKGROUND	5
Table 1: Internal PCGL Acronyms Defining Scope of WG3 Reporting Categories	6
EQUITY STRATIFIER 1: AGE	7
EQUITY STRATIFIER 2: RACE	8
Table 2A: PCGL Race Data Provenance Categories	8
Table 2B: PCGL Racial Reporting Categories.....	11
EQUITY STRATIFIER 3: ETHNIC OR CULTURAL ORIGIN OF PERSON	15
Table 3A: PCGL Ethnicity Data Provenance Categories	15
Table 3B: PCGL Ethnic/Cultural Origin Reporting Categories	17
EQUITY STRATIFIER 4: GENDER	19
Table 4A: PCGL Gender Data Provenance Categories.....	19
Table 4B: PCGL Gender Reporting Categories	21
EQUITY STRATIFIER 5: SEX AT BIRTH	23
Table 5A: PCGL Sex at Birth Data Provenance Categories	23
Table 5B: Sex at Birth Reporting Categories	24
EQUITY STRATIFIER 6: EDUCATION	27
Table 6A: PCGL Education Data Provenance Categories.....	27
Table 6B: Education Reporting Categories	28
EQUITY STRATIFIER 7: PERSONAL INCOME	30
Table 7A: PCGL Income Data Provenance Categories.....	30
Table 7B: Income Reporting Categories	31
EQUITY STRATIFIER 8: GEOGRAPHIC LOCATION.....	33
Table 8A: PCGL Geographic Location Data Provenance Categories.....	33
Table 8B: Geographic Location Reporting Categories	34
REFERENCES	36



REVISION HISTORY

This section/table is necessary for most documents in order to track who has been modifying it, how, and when. This is a living document that will be reviewed and modified in time. The revision history is available below.

Date	Version	Description	Author
2025-03-31	0.1	Created draft with initial content	Alexander Chan
2025-05-05	1.0	Finalized content. Draft ready for publication	Alexander Chan Amirah-Iman Hicks Claude Bhérer



INTRODUCTION

As part of its commitment to equity in genomics, the Pan-Canadian Genome Library (PCGL) is developing standards for collecting, depositing, and sharing individual-level sociodemographic data to support data harmonization and reproducible research on social determinants of health in Canada. This work reflects the 2021 resolution adopted by the 74th World Health Assembly (WHA74.16), which encourages Member States to integrate social determinants of health into public policies and programmes¹. On a national level, this work adheres to the recommendations outlined in the 2022 report, *Pan-Canadian Health Data Strategy: Toward a world-class health data system*², by the Pan-Canadian Health Data Strategy Expert Advisory Group to meet the requirements of the *Canadian Health Data Charter*³. Critically, this report identified systemic fragmentation of Canada's health data as a core barrier preventing health data collection, access, sharing, and use across Canada². We envision that the sociodemographic standards outlined here will support genomics research in meeting the *Canadian Health Data Charter* requirements for standardized terminology, interoperability, and harmonization of person-centric data.

BACKGROUND

The standards outlined in this document were developed with reference to existing Canadian standards – including the Canadian Institute for Health Information (CIHI) and Statistics Canada – and informed by international frameworks outlined within the National Academies of Science, Engineering, and Medicine (NASEM) *Rethinking Race and Ethnicity in Biomedical Research* foundational report⁴, the *All of Us* Research Program⁵, and the UK Biobank⁶.

The following abbreviations specify the intended usage of the standards and reporting categories developed by Working Group 3-Data Diversity and Inclusion (WG3), with the understanding that the PCGL data dictionary developed by Working Group 2-Data Infrastructure, Portals and Tools (WG2) **will need to harmonize sociodemographic data from various exemplar (retrospective) projects while also defining appropriate standards for prospective projects**. Accordingly, **specific reporting categories were created solely for the purpose of harmonizing data from exemplar projects** and would be redundant/excluded by new projects (i.e., prospective projects) consenting participants using PCGL standards. For example, please see the *PCGL Reference Survey for Participant Sociodemographic Data* created by WG3. This reference survey is provided to prospective projects to ensure participant sociodemographic data is aligned with the PCGL standards outlined here.



Table 1: Internal PCGL Acronyms Defining Scope of WG3 Reporting Categories

ACRONYM	TERM	CONSTRUCT
RQ	Reference Question	PCGL developed question provided to the prospective projects ensuring that data across multiple projects follows the PCGL standard for collection, reporting, and sharing.
SV	Study Variable	Internal PCGL usage to store study-specific information (e.g., the date and method of collection for various reporting categories), adhering to various applicable standards (CIHI ⁷ , NASEM ⁴ , etc.) for data provenance.
FD	Field Descriptor	PCGL field value standard, defining the corresponding variable construct, type, and requirements.
PPI	Participant Provided Information	<p>Indicates that data for this reporting category should be asked to participants/self-reported (prospective projects).</p> <p>Where mapping between exemplar projects to PCGL is possible, sociodemographic data from exemplar projects must be harmonized to meet new standards defined by the reporting category.</p>
DH	Data Harmonization	Indicates that data for this reporting category is for internal PCGL usage for data management (e.g., missing values, data entry errors, etc.) and harmonization to exemplar projects before the development of standardized PCGL reporting categories .



EQUITY STRATIFIER 1: AGE

Understanding of social constructs like race, ethnicity and the way participants choose to self-identify may evolve over time. Additionally, even reporting of biological constructs like sex at birth has evolved, as new understandings recognize that sex is influenced not only by external anatomy but also physiology, genetics, and hormones – leading to the increased recognition and reporting of intersex as a valid category. Accordingly, PCGL recommends that the reporting of each sociodemographic variable is associated with an age and date of collection.

PCGL Reference Questions:

Q1. [REQUIRED] What is today's date (YYYY/MM/DD)?

				/			/		
--	--	--	--	---	--	--	---	--	--

Q2. [REQUIRED] What is your year and month of birth (YYYY/MM)?

				/		
--	--	--	--	---	--	--



EQUITY STRATIFIER 2: RACE

In Canada, race is uniquely defined by federal statute through the concept of “visible minority” or “visible minority of person”: a statutory category designated by the 1986 Employment Equity Act as “persons other than Aboriginal persons who are non-Caucasian in race or non-white in colour”^{8–10}. Following this precedent, the PCGL – which aims to address the barriers faced by racialized persons (including Indigenous peoples) in accessing and receiving healthcare – asks participants to self-identify based on the concept of “racialization”. Here, we adopt CIHI’s definition of racialization as a social process “...by which people are judged and categorized into races primarily using differences in physical appearances”¹¹.

We acknowledge that this definition of race – specifically created for a Canadian context – is inherently narrow and reflects certain historical understandings of race that may not capture the full breadth of experience lived by racialized persons in Canada, let alone the rest of the world. Much of this difficulty arises from the fact that race, a social construct, is unfailingly arbitrary and imprecisely defined, historically fluid, as well as politically and culturally dependent^{12,13}; however, without systematic measurement, the effects of social determinants of health (including race) can go unnoticed. This motivated the need to develop a Pan-Canadian standard for race that would capture the full diversity present in Canada. Thus, as part of our work to fulfill the PCGL’s commitment to equity in genomics, we define race as it would be broadly understood socially and evaluated legally, based on sociodemographic standards and legal definitions set by the Government of Canada⁸ and implemented by Statistics Canada in the *Census of Population*^{9,10}, as well as Pan-Canadian health standards set by CIHI^{7,11}.

Table 2A: PCGL Race Data Provenance Categories

VARIABLE	CONSTRUCT	SCOPE
Field Name	race	FD
Field Definition	Race is a social construct used to judge and categorize people based on perceived differences in physical appearance in ways that create and maintain power differentials within social hierarchies. There is no scientifically supported biological basis for discrete racial groups ¹¹ . Participants may have multiple racial identities. All categories selected by the participants should be reported.	FD
Field Requirements	To enable accurate data harmonization within the PCGL, we regard race and ethnicity as distinct social constructs, uniquely defined in Canada by various federal policies, legislation, and Supreme Court rulings (e.g., <i>Employment Equity Act</i> ⁸ , <i>Canadian</i>	FD



	<p><i>Multiculturalism Act</i>¹⁴, <i>Canadian Human Rights Act</i>¹⁵, <i>R. v. Krymowski</i>, [2005] 1 S.C.R. 101, 2005 SCC 7¹⁶, etc.). Accordingly, the “measurement and reporting of racial and ethnic health inequalities should reflect these differences”¹¹, and it is not appropriate to use race, ethnicity, or other proximate variables (e.g., country of birth, country of origin, ancestry, etc.) interchangeably^{17,18}. Similarly, it is not appropriate to aggregate responses (within the PCGL) that measured race using a combination of racial, cultural, ethnic, etc. categories as response options¹⁷.</p> <p>Regarding conformity to other international standards, Statistics Canada states: “There are no comparable internationally recognized standards,” since [visible minority of person] measures a concept that is defined in Canadian legislation,”⁹. Accordingly, particular care must be taken to verify the method of collecting participants’ race before aggregating or comparing results from the PCGL to other international projects.</p>	
Field Type	string	FD
PCGL Reference Question: Race	<p>Q3. [REQUIRED] In our society, people are often described based on race or racial background. There is no scientifically supported biological basis for discrete racial groups. Rather, race is a sociopolitical construct based on perceived differences in physical appearance. This influences the way we are treated by individuals and institutions and may affect our health.</p> <p>With this in mind, which category(ies) best describes you? Select all that apply:</p> <ul style="list-style-type: none"><input type="checkbox"/> Black<input type="checkbox"/> East Asian<input type="checkbox"/> Indigenous (First Nations, Inuk/Inuit, Métis)<input type="checkbox"/> Latin American<input type="checkbox"/> Middle Eastern or North African<input type="checkbox"/> South Asian<input type="checkbox"/> Southeast Asian<input type="checkbox"/> White<input type="checkbox"/> Do not know<input type="checkbox"/> Prefer not to answer<input type="checkbox"/> None of these fully describe me (i.e., another race/racial category)	RQ



	<p><i>Branching Logic: when “None of these fully describe me” selected, then:</i></p> <p>Please specify: <i>(display optional free text)</i></p>	
Race Question	Data submitters are asked to report their question of race, including the provided response options/reporting categories unless following the PCGL reference question and standard reporting categories.	SV
Method of Collection	<p>1. Socially assigned: when a person other than the participant themselves (e.g., a parent, healthcare practitioner, etc.) selected a racial category based on social perceptions of race.</p> <p>2. Self-identified: the participant reported for themselves.</p> <p>3. Derived: race was predicted or imputed (e.g., projecting race from parents, place of birth, genetic ancestry, etc.).</p> <hr/> <p>In line with NASEM standards⁴, data submitters must report how race data was collected by selecting one of the three reporting categories provided above. This is necessary to enable data aggregation, harmonization, and reproducibility. Responses from prospective projects should always be self-identified. To respect participants’ autonomy and right to self-determination, as well as to avoid conflating differences in (genetic) population structure with socially constructed racial categories, responses that were self-identified should not be aggregated with responses that were socially assigned or derived.</p>	SV
Additional Comments	<p>NASEM recognizes that potential bias may exist due to the way racial categories were defined and handled. This may include sampling, classification, method of data collection, as well as the completeness of the data (e.g., response rate)⁴.</p> <hr/> <p>In line with NASEM standards, data submitters may choose to report (via open field text) whether potential bias may exist due to the way racial categories were defined and handled.</p>	SV



Again, we emphasize that race is a socially defined concept – something made especially clear in this context as we draw on legal and sociodemographic frameworks developed by various Canadian institutions to inform the racial categories reported in the following table. They do not reflect any meaningful biological differences and should not be used in place of characteristics (whether biological or sociodemographic) that have not been directly observed or measured.

Table 2B: PCGL Racial Reporting Categories

REPORTING CATEGORY	ADDITIONAL COMMENTS (DEFINITIONS, MAPPING, ETC.)	SCOPE
Black	<p>A person's sense of being Black. Must be declared by the person. Where collected, CIHI considers African, African Canadian, Afro-Caribbean¹¹, as examples of valid identities within this category (which may include others not listed here).</p> <hr/> <p>While a person's place or country of origin is just one of many factors that may contribute to racialization based on perceived differences in physical appearance, we provide relevant examples listed by CIHI¹¹ for clarity.</p>	PPI
East Asian	<p>A person's sense of being East Asian. Must be declared by the person. Where collected, CIHI considers Chinese, Korean, Japanese, Taiwanese¹¹, as examples of valid identities within this category (which may include others not listed here).</p> <hr/> <p>While a person's place or country of origin is just one of many factors that may contribute to racialization based on perceived differences in physical appearance, we provide relevant examples listed by CIHI¹¹ for clarity.</p>	PPI
Indigenous (First Nations, Inuk/Inuit, Métis)	Indigenous identities in North America (i.e., Indigenous peoples identifying as First Nations, Inuk/Inuit, Métis) ¹¹ . Other Indigenous identities are to be reported in the reporting category, "Another Racial Category". This distinction follows current Statistics Canada standards ¹⁹ .	PPI
Latin American	A person's sense of being Latin American. Must be declared by the person.	PPI
Middle Eastern or North African	A person's sense of being Middle Eastern or North African. Must be declared by the person. Where collected, CIHI considers Arab, Persian, West Asian (e.g., Afghan, Egyptian, Iranian, Kurdish,	PPI



	<p>Lebanese, Turkish)¹¹, as examples of valid identities within this category (which may include others not listed here).</p> <hr/> <p>While a person's place or country of origin is just one of many factors that may contribute to racialization based on perceived differences in physical appearance, we provide relevant examples listed by CIHI¹¹ for clarity.</p>	
South Asian	<p>A person's sense of being South Asian. Must be declared by the person. Where collected, CIHI considers Bangladeshi, Indian, Indo-Caribbean, Pakistani, Sri Lankan¹¹, as examples of valid identities within this category (which may include others not listed here).</p> <hr/> <p>While a person's place or country of origin is just one of many factors that may contribute to racialization based on perceived differences in physical appearance, we provide relevant examples listed by CIHI¹¹ for clarity.</p>	PPI
Southeast Asian	<p>A person's sense of being Southeast Asian. Must be declared by the person. Where collected, CIHI considers Cambodian, Filipino, Indonesian, Thai, Vietnamese, as examples of valid identities within this category (which may include others not listed here).</p> <hr/> <p>While a person's place or country of origin is just one of many factors that may contribute to racialization based on perceived differences in physical appearance, we provide relevant examples listed by CIHI¹¹ for clarity.</p>	PPI
White	<p>A person's sense of being White. Must be declared by the person.</p>	PPI
Another Racial Category	<p>A response which does not fall into any of the other permissible racial reporting categories (e.g., Pacific Islander, Indigenous identities outside of North America, etc.).</p> <hr/> <p>A response of "None of these fully describe me" from prospective projects. A response of "Other" from exemplar projects. A racial reporting category which does not fall into any of the previous reporting categories (e.g., Pacific Islander).</p>	PPI
Multiple Racial Identities	<p>A response with more than one of the above categories.</p> <hr/> <p>Internal PCGL data processing ONLY (Prospective projects should instead allow participants to select multiple racial</p>	DH



	<p>reporting categories). Adopting similar methodology employed by Statistics Canada to prevent multiple counts from a single individual (e.g., a respondent who self-identifies as “Black” and “East Asian” would not be counted both in the “Black” and “East Asian” reporting categories)¹⁰. Also equivalent to the <i>All of Us</i> “More than one category” used in the data snapshot “Self-reported Categories” graph:</p> <div><div><div>White</div><div>American Indian or Alaska Native</div><div>Asian</div><div>Black, African American or African</div><div>Hispanic Latino or Spanish</div><div>More than one category</div><div>Other</div><div>Prefer not to say</div></div><table><thead><tr><th>Category</th><th>Percentage</th></tr></thead><tbody><tr><td>White</td><td>57.7%</td></tr><tr><td>American Indian or Alaska Native</td><td>4.1%</td></tr><tr><td>Asian</td><td>4.2%</td></tr><tr><td>Black, African American or African</td><td>18.7%</td></tr><tr><td>Hispanic Latino or Spanish</td><td>19.6%</td></tr><tr><td>More than one category</td><td>7.0%</td></tr><tr><td>Other</td><td>2.4%</td></tr><tr><td>Prefer not to say</td><td>0.5%</td></tr></tbody></table></div>	Category	Percentage	White	57.7%	American Indian or Alaska Native	4.1%	Asian	4.2%	Black, African American or African	18.7%	Hispanic Latino or Spanish	19.6%	More than one category	7.0%	Other	2.4%	Prefer not to say	0.5%	
Category	Percentage																			
White	57.7%																			
American Indian or Alaska Native	4.1%																			
Asian	4.2%																			
Black, African American or African	18.7%																			
Hispanic Latino or Spanish	19.6%																			
More than one category	7.0%																			
Other	2.4%																			
Prefer not to say	0.5%																			
Do not know	<p>When the question of race was asked, but the answer is unknown.</p> <hr/> <p>A response of “Do not know” from prospective projects. A response of “Unknown” from exemplar projects (if this was a permissible response option, otherwise see reporting category: Missing – Unknown).</p>	PPI																		
Prefer not to answer	<p>Participant chooses not to or does not feel comfortable reporting.</p> <hr/> <p>A response of “Prefer not to answer” from prospective projects.</p>	PPI																		
Missing – Unknown	<p>Information is not stated but the question was asked (i.e., a value is applicable but unknown). For example, the question of race was asked, but was skipped/not answered (left blank). This is distinct from a participant response of “Do not know” or “Unknown”, in which case a valid response is provided and is not “missing”.</p>	DH																		
Missing – Not collected	<p>Information was not recorded/not asked by the project or study.</p>	DH																		
Missing – Not provided	<p>Information of an expected format was not given (data entry error); a value may be given at a later stage.</p>	DH																		



Missing – Restricted access	Information exists (question was asked) but cannot be released openly because of privacy concerns.	DH
Date of Sociodemographic Survey Completion	<p>Race is a social construct, and the way participants self-identify may evolve over time.</p> <hr/> <p>Following NASEM recommendations⁴, responses should include a date (e.g., the date when the participant completed the study survey) in the metadata.</p>	SV



EQUITY STRATIFIER 3: ETHNIC OR CULTURAL ORIGIN OF PERSON

Although this has not always been the case, ethnicity (i.e., the ethnic or cultural origin of person) is currently recognized in Canada as a sociodemographic construct distinct from race. For instance, in the 2021 *Census of Population*, respondents were asked to self-identify as a visible minority (or not) and, separately, were also asked to self-identify the ethnic or cultural origins of their ancestors²⁰. In the context of the PCGL, we adopt CIHI's definition of ethnicity (see Table 3A), which – unlike race – is not explicitly defined by Canadian law.

Like race, ethnicity is a socially defined concept. Similarly, ethnic groups do not reflect any meaningful biological differences and should not be used in place of characteristics (whether biological or sociodemographic) that have not been directly observed or measured.

Table 3A: PCGL Ethnicity Data Provenance Categories

VARIABLE	CONSTRUCT	SCOPE
Field Name	ethnicity	FD
Field Definition	Ethnicity is a “multi-dimensional concept referring to community belonging and a shared cultural group membership. It is related to sociodemographic characteristics, including language, religion, geographic origin, nationality, cultural traditions, ancestry and migration history, among others” ¹¹ .	FD
Field Requirements	To enable accurate data harmonization within the PCGL, we regard race and ethnicity as distinct social constructs, uniquely defined in Canada by various federal policies, legislation, and Supreme Court rulings (e.g., <i>Employment Equity Act</i> ⁸ , <i>Canadian Multiculturalism Act</i> ¹⁴ , <i>Canadian Human Rights Act</i> ¹⁵ , <i>R. v. Krymowski</i> , [2005] 1 S.C.R. 101, 2005 SCC 7 ¹⁶ , etc.). Accordingly, the “measurement and reporting of racial and ethnic health inequalities should reflect these differences” ¹¹ , and it is not appropriate to use race, ethnicity, or other proximate variables (e.g., country of birth, country of origin, ancestry, etc.) interchangeably ^{17,18} . Similarly, it is not appropriate to aggregate responses (within the PCGL) that measured race using a combination of racial, cultural, ethnic, etc. categories as response options ¹⁷ .	FD
Field Type	string	FD



PCGL Reference Question: Ethnic or Cultural Origin of Person	<p>Q4. [REQUIRED] Ethnic or cultural origin refers to the multi-dimensional concept of community belonging and a shared cultural group membership. It can include language, religion, country of origin, nationality, cultural traditions, and migration history of a person's ancestors (biological parents, grandparents, etc.). For examples of ethnic or cultural origins, visit Statistics Canada's Examples of ethnic or cultural origins web page.</p> <p>What were the ethnic or cultural origins of your ancestors? Specify as many origins as applicable using capital letters: [Display free text]</p> <p>Or,</p> <p><input type="checkbox"/> Do not know <input type="checkbox"/> Prefer not to answer</p>	RQ
Ethnic or Cultural Origin of Person Question	Data submitters are asked to report their question of ethnicity, including the provided response options/reporting categories unless following the PCGL reference question and standard reporting categories.	SV
Method of Collection	<ol style="list-style-type: none">Socially assigned: when a person other than the participant themselves specified an ethnic or cultural origin based on social perceptions of ethnicity/cultural origins.Self-identified: the participant reported for themselves.Derived: ethnicity was predicted or imputed (e.g., projecting ethnicity from parents, place of birth, genetic ancestry, etc.). <hr/> <p>In line with NASEM standards⁴, data submitters must report how ethnicity data was collected by selecting one of the three reporting categories provided above. This is necessary to enable data aggregation, harmonization, and reproducibility. Responses from prospective projects should always be self-identified. To respect participants' autonomy and right to self-determination, as well as to avoid conflating differences in (genetic) population structure with socially constructed ethnic groups, responses that were self-identified should not be aggregated with responses that were socially assigned or derived.</p>	SV
Additional Comments	Like race, bias may exist due to the way ethnic/cultural groups were defined and handled. This may include sampling, classification, method of data collection, as well as the completeness of the data (e.g., response rate).	SV



	Data submitters may choose to report (via open field text) whether potential bias may exist due to the way ethnic groups were defined and handled.	
--	--	--

Table 3B: PCGL Ethnic/Cultural Origin Reporting Categories

REPORTING CATEGORY	ADDITIONAL COMMENTS (DEFINITIONS, MAPPING, ETC.)	SCOPE
<i>Free Text Input</i>	<p>Following CIHI recommendations¹¹, ethnic and cultural origin should mimic Statistics Canada standards¹⁹ with an open field text and the link to the 2021 Statistics Canada <i>Examples of ethnic or cultural origins</i>.</p> <p>PCGL largely adopts CIHI's recommendations for permissible response options to capture the most granular demographic information and enable harmonization with Statistics Canada. Minor modifications were made to provide participants with the option to decline to answer. Furthermore, while we recognize that the Statistics Canada <i>Examples of ethnic or cultural origins</i> is updated periodically (with the newest version last modified on April 4, 2024 at the time of this release), we ask prospective projects to continue using the 2021 version. This enables the PCGL to benchmark against a stable version and allow harmonization and reproducibility.</p>	PPI
Another Ethnic or Cultural Origin	A response of "Other" from retrospective exemplar projects. Participants in prospective projects should instead be allowed to freely self-identify with an open text field.	DH
Multiple Ethnic/Cultural Origins	<p>A response with more than one of the above categories.</p> <p>Internal PCGL data processing ONLY (prospective projects should instead allow participants to list multiple ethnic and/or cultural origins). Adopting similar methodology employed by Statistics Canada to prevent multiple counts from a single individual¹⁹.</p>	DH
Do not know	<p>When the question of ethnicity was asked, but the answer is unknown.</p> <p>A response of "Do not know" from prospective projects. A response of "Unknown" from exemplar projects (if this was a</p>	PPI



	permissible response option, otherwise see reporting category: Missing – Unknown).	
Prefer not to answer	Participant chooses not to or does not feel comfortable reporting. <hr/> A response of “Prefer not to answer” from prospective projects.	PPI
Not applicable	Information is inappropriate to ask (question should be excluded from study).	DH
Missing – Unknown	Information is not stated but the question was asked (i.e., a value is applicable but unknown). For example, the question of ethnicity was asked, but was skipped/not answered (left blank). This is distinct from a participant response of “Do not know” or “Unknown”, in which case a valid response is provided and is not “missing”.	DH
Missing – Not collected	Information was not recorded/not asked by the project or study.	DH
Missing – Not provided	Information of an expected format was not given (data entry error); a value may be given at a later stage.	DH
Missing – Restricted access	Information exists but cannot be released openly because of privacy concerns.	DH
Date of Sociodemographic Survey Collection	Ethnicity is a social construct, and the way participants self-identify may evolve over time. <hr/> Responses should include a date (e.g., the date when participant completed study survey) in the metadata.	SV



EQUITY STRATIFIER 4: GENDER

Within the PCGL, we largely adopt CIHI's Pan-Canadian standards for defining gender (see Table 4A), which were informed by a retrospective review of gender-related data collected by federal administrative systems, provincial and territorial health agencies, and electronic medical record providers across the Canadian healthcare system^{7,21}.

Table 4A: PCGL Gender Data Provenance Categories

VARIABLE	CONSTRUCT	SCOPE
Field Name	gender identity	FD
Field Definition	Gender refers to the socially constructed roles, behaviours, expressions and identities of girls, women, boys, men, and gender diverse people ²² . A person's gender may be the same as or different from their assigned sex at birth and may change over time ^{7,23} .	FD
Field Requirements	<p>Response options that acknowledge categories beyond the gender binary must be provided. See Table 4B for examples which meet the PCGL's minimum reporting requirements.</p> <p>When response options were limited to the gender binary (e.g., "woman", "man", "unknown/not applicable"), the resulting data is not sufficient to report on gender identity. This data should be labelled and reported on as "Recorded sex or gender" and should not be aggregated with other data which meets the minimum reporting requirements^{7,23}.</p>	FD
Field Type	string	FD
PCGL Reference Question: Gender	<p>Q6. [REQUIRED] Gender refers to the socially constructed roles, behaviours, expressions and identities of girls, women, boys, men, and gender diverse people. A person's gender may be the same as or different from their assigned sex at birth and may change over time.</p> <p>Which category best expresses how you currently describe your gender identity?</p> <p><input type="checkbox"/> Man</p> <p><input type="checkbox"/> Woman</p> <p><input type="checkbox"/> Another Gender</p>	RQ



	<p><i>Branching Logic: when “Another Gender” selected, then:</i></p> <p>Please specify: (for example, Non-binary, Two-spirit, Trans, etc.) <i>(display optional free text)</i></p> <p><input type="checkbox"/> Prefer not to answer <input type="checkbox"/> Not applicable</p> <p><i>Response option provided when the question of gender is not appropriate to ask (e.g., newborn studies). This response option should NOT be provided to adult participants.</i></p>	
Gender Question	Data submitters are asked to report their question of gender, including the provided response options/reporting categories unless following the PCGL reference question and standard reporting categories.	SV
Method of Collection	<p>1. Self-identified: the participant reported for themselves. 2. Other: other methods of collection, which may include health card registration and other administrative sources following the 2017 amendment of the <i>Canadian Human Rights Act</i>²³.</p> <hr/> <p>Data submitters must report how gender data was collected by selecting one of the two reporting categories provided above. This is necessary to enable data aggregation, harmonization, and reproducibility. Responses from prospective projects should always be self-identified. To respect participants’ autonomy and right to self-determination, responses that were self-identified should not be aggregated with other methods of collection, particularly if responses beyond the gender binary were not provided.</p> <p>For data that is sourced from health card or other administrative data, PCGL adopts CIHI standards, which recognizes that this may depend on whether an individual has updated their official or legal documentation. Accordingly this data should be relabelled or described as “Recorded sex or gender,”^{7,23}.</p>	SV



Table 4B: PCGL Gender Reporting Categories

REPORTING CATEGORY	ADDITIONAL COMMENTS (DEFINITIONS, MAPPING, ETC.)	SCOPE
Man	<p>A person's sense of being a man. Must be declared by the person. Where collected, male/boy are also valid identities within this category²³.</p> <hr/> <p>CIHI standards uses reporting categories <i>male</i> and <i>female</i> for both gender identity and sex at birth, which “facilitates easier cross-tabulation to determine whether an individual’s current gender identity or does not match their assigned sex at birth,”²³. Additionally, “CIHI believes the approach of using <i>female</i> and <i>male</i> as reporting categories is clearer than man/woman/boy/girl because it applies across all age groups and is supported in the literature,”²³. However, to avoid conflating gender (a social construct) with assigned sex at birth, PCGL instead recommends using <i>man</i> and <i>woman</i> as reporting categories.</p>	PPI
Woman	<p>A person's sense of being a woman. Must be declared by the person. Where collected, female/girl are also valid identities within this category²³.</p> <hr/> <p>CIHI standards uses reporting categories <i>male</i> and <i>female</i> for both gender identity and sex at birth, which “facilitates easier cross-tabulation to determine whether an individual’s current gender identity or does not match their assigned sex at birth,”²³. Additionally, “CIHI believes the approach of using <i>female</i> and <i>male</i> as reporting categories is clearer than man/woman/boy/girl because it applies across all age groups and is supported in the literature,”²³. However, to avoid conflating gender (a social construct) with assigned sex at birth, PCGL instead recommends using <i>man</i> and <i>woman</i> as reporting categories.</p>	PPI
Another Gender	<p>A response which does not fall along the traditional binary gender spectrum. Includes persons who report being unsure of their gender, being both male and female, or being neither male nor female (e.g., a non-binary gender)²³.</p> <hr/> <p>Following CIHI guidance, an optional free text subcategory is suggested to collect and report more granular data (e.g., non-binary, two-spirit) where sample sizes are sufficient.</p>	PPI



Prefer not to answer	Participant chooses not to or does not feel comfortable reporting.	PPI
Not applicable	Information is inappropriate to ask (question should be excluded from study).	DH
Do not know	<p>A response of “Unknown” from exemplar projects (if this was a permissible response option, otherwise see reporting category: Missing – Unknown).</p> <hr/> <p>Aligning with CIHI standard reporting categories²³, PCGL does not suggest “Do not know” as a valid response option. It is provided here for data harmonization purposes only.</p>	DH
Missing – Unknown	Information is not stated but the question was asked (i.e., a value is applicable but unknown). For example, the question of gender was asked, but was skipped/not answered (left blank). This is distinct from a participant response of “Do not know” or “Unknown”, in which case a valid response is provided and is not “missing”.	DH
Missing – Not collected	Information was not recorded/not asked by the project or study.	DH
Missing – Not provided	Information of an expected format was not given (data entry error); a value may be given at a later stage.	DH
Missing – Restricted access	Information exists but cannot be released openly because of privacy concerns.	DH
Date of Sociodemographic Survey Collection	<p>Gender is a social construct, and the way participants self-identify may evolve over time.</p> <hr/> <p>Responses should include a date (e.g., the date when participant completed study survey) in the metadata.</p>	SV



EQUITY STRATIFIER 5: SEX AT BIRTH

A person's assigned sex at birth is a biological concept historically defined by external anatomy and subsequently viewed as a binary attribute. Increasingly, sex at birth has been assigned by medical professionals based on an individual's "chromosomes, hormone levels, sex organs, and secondary sex characteristics,"²⁴. However, because genetics is often used to misconceptualize sex as binary (i.e., male XY or female XX), it is important to emphasize that "several other chromosomal arrangements, variations in gene expression patterns, and hormone levels exist,"²⁴.

Table 5A: PCGL Sex at Birth Data Provenance Categories

VARIABLE	CONSTRUCT	SCOPE
Field Name	assigned sex at birth	FD
Field Definition	Sex is a biological concept that has been historically defined solely based on external anatomy; however, sex is also influenced by our physiology, genes, and hormones ⁷ .	FD
Field Requirements	Assignment of sex at birth is typically based on external anatomy and limited to female and male. Yet, sex is a complex biological concept that includes anatomy, physiology, genes, and hormones ⁷ . Increasingly, sex at birth may also be assigned as intersex, which is a term used to describe a range of sex characteristics. Not all intersex variations are visible or detected at birth, but when they are, sex at birth may be recorded as "intersex" or assigned to "female" or "male". Accordingly, additional consultation may be needed to report on this category, as many individuals identify as intersex later in life ²³ .	FD
Field Type	string	FD
PCGL Reference Question: Sex at Birth	Q8. [REQUIRED] What was your biological sex assigned at birth? <input type="checkbox"/> Male <input type="checkbox"/> Female <input type="checkbox"/> Intersex <input type="checkbox"/> None of these describe me <i>Branching Logic: when "None of these describe me" selected, then:</i>	RQ



	<p>Please specify: <i>(display optional free text; can include clinician comments if relevant)</i></p> <p><input type="checkbox"/> Prefer not to answer</p>	
Sex at Birth Question	Data submitters are asked to report their question of assigned sex at birth, including the provided response options/reporting categories unless following the PCGL reference question and standard reporting categories.	SV
Method of Collection	<p>1. Self-identified: the participant reported for themselves.</p> <p>2. Clinician-recorded: assigned by a qualified healthcare provider where clinically relevant.</p> <p>3. Derived: sex at birth determined through genetic analysis.</p> <p>4. Other: other methods of collection, including health cards or other administrative sources.</p> <hr/> <p>Data submitters must report how biological sex data was collected by selecting one of the three reporting categories provided above. This is necessary to enable data aggregation, harmonization, and reproducibility. Responses from prospective projects should always be self-identified or clinician-recorded (where appropriate).</p> <p>For data that is sourced from health card or other administrative data, PCGL adopts CIHI standards, which recognizes that this may depend on whether an individual has updated their official or legal documentation. Accordingly this data should be relabelled or described as “Recorded sex or gender,”^{7,23}. This data should not be aggregated with self-identified or clinician-recorded data.</p>	SV

Table 5B: Sex at Birth Reporting Categories

REPORTING CATEGORY	ADDITIONAL COMMENTS (DEFINITIONS, MAPPING, ETC.)	SCOPE
Male	Sex assigned and recorded at a person's birth (e.g., recorded on the original birth certificate).	PPI
Female	Sex assigned and recorded at a person's birth (e.g., recorded on the original birth certificate).	PPI



Intersex	<p>Includes any response that may refer to individuals with sex characteristics that do not conform to binary notions of male or female bodies, including disorders/diverse/differences of sex development or variations of sex characteristics²⁵.</p> <hr/> <p>Additional consultation may be needed to report on this category, as many individuals identify as intersex later in life⁷.</p>	PPI
Another Sex	<p>A response which does not fall into any of the previous reporting categories.</p> <hr/> <p>A response of “None of these fully describe me” from prospective projects. An optional free text subcategory is suggested to collect more granular data, including clinician comments where relevant (e.g., description of secondary sex characteristics).</p>	PPI
Prefer not to answer	Participant chooses not to or does not feel comfortable reporting.	PPI
Do not know	A response of “Unknown” from exemplar projects.	DH
Missing – Unknown	Information is not stated but the question was asked (i.e., a value is applicable but unknown). For example, the question of sex assigned at birth was asked, but was skipped/not answered (left blank). This is distinct from a participant response of “Do not know” or “Unknown”, in which case a valid response is provided and is not “missing”.	DH
Missing – Not collected	Information was not recorded/not asked by the project or study.	DH
Missing – Not provided	Information of an expected format was not given (data entry error); a value may be given at a later stage.	DH
Missing – Restricted access	Information exists but cannot be released openly because of privacy concerns.	DH
Date of Sociodemographic Survey Collection	Although sex at birth is NOT a social construct, sex at birth has rarely been recorded as intersex and remains underreported (though this practice is increasing over time) ²⁵ . Consequently, the way participants self-identify may evolve over time, as many intersex traits do not become apparent until puberty or later in life ^{7,24} .	SV



	Responses should include a date (e.g., the date when participant completed study survey) in the metadata.	
--	---	--



EQUITY STRATIFIER 6: EDUCATION

PCGL uses educational attainment to measure and report education-related inequality, adopting CIHI standards for education-related inequality in the population. In this context, educational attainment refers to the highest level of schooling achieved⁷.

Table 6A: PCGL Education Data Provenance Categories

VARIABLE	CONSTRUCT	SCOPE
Field Name	education	FD
Field Definition	The highest level of education attained.	FD
Field Requirements	PCGL largely adopts the Statistics Canada reporting categories for education, with minor modifications to capture increased granularity and provide participants with the option to choose not to answer.	FD
Field Type	string	FD
PCGL Reference Question: Education	<p>Q10. [REQUIRED] What is the highest level of schooling you attained? Include qualifications obtained in Canada or outside Canada.</p> <ul style="list-style-type: none"><input type="checkbox"/> No formal education<input type="checkbox"/> Elementary school or equivalent<input type="checkbox"/> High school (secondary school) diploma or equivalency certificate (for example, General Education Development, or Adult Basic Education)<input type="checkbox"/> Certificate of Apprenticeship, Certificate of Qualification (Journey person's designation), or other trades certificate or diploma (for example, hairstyling, cooking, electrician, carpentry)<input type="checkbox"/> College, CEGEP, or other non-university certificate or diploma (for example, accounting technology, industrial engineering technology, legal assistant)<input type="checkbox"/> Bachelor's degree (for example, B.A., B.A.(Hons.), B.Sc., B.Ed., LL.B.)<input type="checkbox"/> Degree in medicine, dentistry, veterinary medicine or optometry (M.D., D.D.S., D.M.D., D.V.M., O.D.)<input type="checkbox"/> Master's degree (for example, M.A., M.Sc., M.Ed., M.B.A.)<input type="checkbox"/> Doctoral degree (for example, Ph.D.)<input type="checkbox"/> Post-doctoral fellowship or training<input type="checkbox"/> Prefer not to answer	RQ



	<input type="checkbox"/> Not applicable	
Education Question	Data submitters are asked to report their question of education, including the provided response options/reporting categories unless following the PCGL reference question and standard reporting categories.	SV
Method of Collection	1. Self-identified: the participant reported for themselves. 2. Derived: a study which mapped a response to the PCGL reporting categories.	SV
	Data submitters must report how education data was collected by selecting one of the two reporting categories provided above.	

Table 6B: Education Reporting Categories

REPORTING CATEGORY	ADDITIONAL COMMENTS (DEFINITIONS, MAPPING, ETC.)	SCOPE
No formal education	-	PPI
Elementary school or equivalent	-	PPI
High school (secondary school) diploma or equivalency certificate	Among others, may include General Education Development, or Adult Basic Education.	PPI
Certificate of Apprenticeship, Certificate of Qualification, or other trades certificate or diploma	Among others, may include Journeyperson's designation, or trades certificate or diploma in hairstyling, cooking, electrician, carpentry, etc.	PPI
College, CEGEP, or other non-university certificate or diploma	-	PPI
Bachelor's degree	Among others, may include B.A., B.A. (Hons)., B.Sc., B.Ed., LL.B.	PPI
Degree in medicine, dentistry, veterinary medicine or optometry	Among others, may include M.D., D.D.S., D.M.D., D.V.N., O.D.	PPI



Master's degree	Among others, may include M.A., M.Sc., M.Ed., M.B.A.	PPI
Doctoral degree	Among others, may include PhD, DPhil.	PPI
Post-doctoral fellowship or training	-	PPI
Prefer not to answer	Participant chooses not to or does not feel comfortable reporting.	PPI
Not applicable	Information is inappropriate to ask (question should be excluded from study).	DH
Missing – Unknown	Information is not stated but the question was asked (i.e., a value is applicable but unknown). This is distinct from a participant response of “Do not know” or “Unknown”, in which case a valid response is provided and is not “missing”.	DH
Missing – Not collected	Information was not recorded/not asked by the project or study.	DH
Missing – Not provided	Information of an expected format was not given (data entry error); a value may be given at a later stage.	DH
Missing – Restricted access	Information exists but cannot be released openly because of privacy concerns.	DH
Date of Sociodemographic Survey Collection	Responses should include a date (e.g., the date when participant completed study survey) in the metadata.	SV



EQUITY STRATIFIER 7: PERSONAL INCOME

PCGL asks for a person's **personal** income **after** taxes to measure and report income-related inequality in the population. In combination with a participant's geographic location (i.e., their address, defined in the following Equity Stratifier 8), a person's relative income – their income compared with that of others in society⁷ – can be reported to capture differences in socioeconomic status and purchasing power.

Projects that are interested may collect both personal and household-level income (as done in Statistics Canada); however, the PCGL recommends that projects minimally report personal income. The income categories reporting categories provided in the reference question (see Table 7A) were developed in reference to Statistics Canada's *Income Explorer, 2021 Census*²⁶. Although the original percentiles (from Statistics Canada) are stratified by age group, the PCGL applies a unified binning scheme that roughly approximates the 25th through 99th percentiles of after-tax income across the general population. This approach enforces consistency and enables comparability across studies (regardless of the age of recruited participants) with rough alignment to national income distributions.

Table 7A: PCGL Income Data Provenance Categories

VARIABLE	CONSTRUCT	SCOPE
Field Name	personal income	FD
Field Definition	Personal annual income from all sources after taxes.	FD
Field Requirements	-	FD
Field Type	-	FD
PCGL Reference Question: Income	Q11. [REQUIRED] What is your current personal income from all sources <i>after</i> taxes last year? <input type="checkbox"/> Less than \$15,000 <input type="checkbox"/> \$ 15,000 - \$ 19,999 <input type="checkbox"/> \$ 20,000 - \$ 29,000 <input type="checkbox"/> \$ 30,000 - \$ 49,999 <input type="checkbox"/> \$ 50,000 - \$ 69,999 <input type="checkbox"/> \$ 70,000 - \$ 84,999 <input type="checkbox"/> \$ 85,000 - \$ 99,999 <input type="checkbox"/> \$ 100,000 - \$ 124,999 <input type="checkbox"/> \$ 125,000 - \$ 149,999 <input type="checkbox"/> \$ 150,000 or more	RQ



	<input type="checkbox"/> Prefer not to answer <input type="checkbox"/> Not applicable	
Income Question	Data submitters are asked to report their question of income, including the provided response options/reporting categories unless following the PCGL reference question and standard reporting categories.	SV
Method of Collection	1. Self-identified: the participant reported for themselves. 2. Derived: a study which mapped a response to the PCGL reporting categories. <hr/> Data submitters must report how income data was collected by selecting one of the two reporting categories provided above.	SV

Table 7B: Income Reporting Categories

REPORTING CATEGORY	ADDITIONAL COMMENTS (DEFINITIONS, MAPPING, ETC.)	SCOPE
Less than \$ 15,000	-	PPI
\$ 15,000 - \$ 19,999	-	PPI
\$ 20,000 - \$ 29,000	-	PPI
\$30,000 - \$ 49,999	-	PPI
\$ 50,000 - \$ 69,999	-	PPI
\$ 70,000 - \$ 84,999	-	PPI
\$ 85,000 - \$ 99,999	-	PPI
\$ 100,000 - \$ 124,999	-	PPI
\$ 125,000 - \$ 149,999	-	PPI
\$ 150,000 or more	-	PPI
Prefer not to answer	Participant chooses not to or does not feel comfortable reporting.	PPI



Not applicable	Information is inappropriate to ask (question should be excluded from study).	DH
Missing – Unknown	Information is not stated but the question was asked (i.e., a value is applicable but unknown). This is distinct from a participant response of “Do not know” or “Unknown”, in which case a valid response is provided and is not “missing”.	DH
Missing – Not collected	Information was not recorded/not asked by the project or study.	DH
Missing – Not provided	Information of an expected format was not given (data entry error); a value may be given at a later stage.	DH
Missing – Restricted access	Information exists but cannot be released openly because of privacy concerns.	DH
Date of Sociodemographic Survey Collection	Responses should include a date (e.g., the date when participant completed study survey) in the metadata.	SV



EQUITY STRATIFIER 8: GEOGRAPHIC LOCATION

PCGL uses a participant's postal code in conjunction with Statistic Canada's *Postal Code Conversion File*²⁷ (PCCF) to link individual-level data (e.g., health, income, education, etc.) to area-level (neighbourhood) data. Additionally, in line with CIHI standards, a participant's postal code can be used to measure and report on health accessibility in the population based on the categorization of a participant's address as urban or rural/remote using the PCCF. Following CIHI standards, the PCGL considers Statistical Area Classifications (SAC)²⁸ 1-3 as urban areas, and SAC 4-8 as rural/remote⁷.

Table 8A: PCGL Geographic Location Data Provenance Categories

VARIABLE	CONSTRUCT	SCOPE										
Field Name	address	FD										
Field Definition	A participant’s postal code at their current address of residence. When postal code is less specific, additional information may be reported.	FD										
Field Requirements	While only a participant’s postal code is required to link individual-level data with the PCCF, participants are also asked for their city, and province or territory, to help clarify addresses in certain instances (e.g., institutional residences, First Nations lands, and Inuit land claims agreements).	FD										
Field Type	-	FD										
PCGL Reference Question: Geographic Location	<div><div><div>Q13. [REQUIRED] What is your current address of residence? While your postal code is the main information we use, we also ask for your city, and province/territory to help clarify your address in certain instances where postal codes may be less specific – for example, institutional residences, First Nations lands, and Inuit land claim agreements.</div><table><tr><td>Address Line 1 (optional)</td><td></td></tr><tr><td>Address Line 2 (optional)</td><td></td></tr><tr><td>City</td><td></td></tr><tr><td>Province/Territory</td><td></td></tr><tr><td>Postal Code</td><td></td></tr></table></div></div>	Address Line 1 (optional)		Address Line 2 (optional)		City		Province/Territory		Postal Code		RQ
Address Line 1 (optional)												
Address Line 2 (optional)												
City												
Province/Territory												
Postal Code												



	Country (<i>optional</i>)	
Geographic Location Question	Data submitters are asked to report their question of geographic location, including the provided response options/reporting categories unless following the PCGL reference question and standard reporting categories.	SV
Method of Collection	<ol style="list-style-type: none"> Self-identified: the participant reported for themselves. Derived: a study which mapped a response to the PCGL reporting categories. <hr/> Data submitters must report how geographic location data was collected by selecting one of the two reporting categories provided above.	SV

Table 8B: Geographic Location Reporting Categories

REPORTING CATEGORY	ADDITIONAL COMMENTS (DEFINITIONS, MAPPING, ETC.)	SCOPE
Address Line 1	Optional	PPI
Address Line 2	Optional	
City	Required	PPI
Province/Territory	Required	PPI
Postal Code	Required	PPI
Country	Optional	PPI
Missing – Unknown	Information is not stated but the question was asked (i.e., a value is applicable but unknown). This is distinct from a participant response of “Do not know” or “Unknown”, in which case a valid response is provided and is not “missing”.	DH
Missing – Not collected	Information was not recorded/not asked by the project or study.	DH
Missing – Not provided	Information of an expected format was not given (data entry error); a value may be given at a later stage.	DH



Missing – Restricted access	Information exists but cannot be released openly because of privacy concerns.	DH
Date of Sociodemographic Survey Collection	Responses should include a date (e.g., the date when participant completed study survey) in the metadata.	SV



REFERENCES

1. World Health Organization. Operational framework for monitoring social determinants of health equity. (2024).
2. Vivek Goel *et al.* *Pan-Canadian Health Data Strategy: Toward a World-Class Health Data System*. <https://www.canada.ca/content/dam/phac-aspc/documents/corporate/mandate/about-agency/external-advisory-bodies/list/pan-canadian-health-data-strategy-reports-summaries/expert-advisory-group-report-03-toward-world-class-health-data-system/expert-advisory-group-report-03-toward-world-class-health-data-system.pdf> (2022).
3. Health Canada. Pan-Canadian Health Data Charter. *Government of Canada* <https://www.canada.ca/en/health-canada/corporate/transparency/health-agreements/shared-health-priorities/working-together-bilateral-agreements/pan-canadian-data-charter.html> (2023).
4. Committee on the Use of Race and Ethnicity in Biomedical Research *et al.* *Rethinking Race and Ethnicity in Biomedical Research*. 27913 (National Academies Press, Washington, D.C., 2024). doi:10.17226/27913.
5. All of Us Research Program. Survey Explorer. *All of Us Research Hub* <https://www.researchallofus.org/data-tools/survey-explorer/> (2023).
6. UK Biobank. Questionnaire Data. *About our Data* <https://www.ukbiobank.ac.uk/enable-your-research/about-our-data/questionnaire-data> (2019).
7. Flores, F. *et al.* CIHI Reference Data Model Toolkit. **5.3**, (2024).
8. *Employment Equity Act* (S.C. 1995, c. 44). (1995).
9. Statistics Canada. Visible minority of person. *Statistics Canada* <https://www23.statcan.gc.ca/imdb/p3Var.pl?Function=DECI&Id=257515> (2015).
10. Statistics Canada. Table 98-10-0352-01 Visible minority by gender and age: Census divisions. <https://doi.org/10.25318/9810035201-ENG> (2023).
11. Canadian Institute for Health Information. *Guidance on the Use of Standards for Race-Based and Indigenous Identity Data Collection and Health Reporting in Canada*. (2022).
12. Boyd, M. Ethnicity and Race in Canada: Past and Present. in *The International Handbook of the Demography of Race and Ethnicity* (eds. Sáenz, R., Embrick, D. G. & Rodríguez, N. P.) 23–45 (Springer Netherlands, Dordrecht, 2015). doi:10.1007/978-90-481-8891-8_2.



13. Brown, A. The changing categories the U.S. census has used to measure race. *Pew Research Center* <https://www.pewresearch.org/short-reads/2020/02/25/the-changing-categories-the-u-s-has-used-to-measure-race/> (2020).
14. *Canadian Multiculturalism Act (R.S.C., 1985, c. 24)*. (1988).
15. *Canadian Human Rights Act (R.S.C., 1985, c. H-6)*. (1977).
16. *R. v. Krymowski*, [2005] 1 S.C.R. 101, 2005 SCC 7. <https://decisions.scc-csc.ca/scc-csc/scc-csc/en/item/2206/index.do> (2005).
17. Fiona Clement et al. *Guide on Using Categorical Race and Ethnicity Variables*. https://www2.gov.bc.ca/assets/gov/british-columbians-our-governments/multiculturalism-anti-racism/anti-racism/anti-racism-hub/anti-racism-stats-and-research/arda_categoricalraceethnicityvariables2023.pdf (2023).
18. Committee on the Use of Race, Ethnicity, and Ancestry as Population Descriptors in Genomics Research et al. *Using Population Descriptors in Genetics and Genomics Research: A New Framework for an Evolving Field*. 26902 (National Academies Press, Washington, D.C., 2023). doi:10.17226/26902.
19. Statistics Canada. Table 98-10-0341-01 Religion by ethnic or cultural origins: Canada, provinces and territories and census metropolitan areas with parts. <https://doi.org/10.25318/9810034101-eng> (2023).
20. Statistics Canada. Archived - 2021 Census: 2A-L. https://www.statcan.gc.ca/en/statistical-programs/instrument/3901_Q2_V6 (2020).
21. Shannon O'Connor & Dana Riley. CIHI's Equity Stratifiers: Modernizing our sex and gender data standards. (2023).
22. Canadian Institutes of Health Research. What is gender? What is sex? *Government of Canada* <https://cihr-irsc.gc.ca/e/48642.html> (2014).
23. Canadian Institute for Health Information. *In Pursuit of Health Equity: Defining Stratifiers for Measuring Health Inequality — A Focus on Age, Sex, Gender, Income, Education and Geographic Location*. (2018).
24. King, D. E. The Inclusion of Sex and Gender Beyond the Binary in Toxicology. *Front Toxicol* **4**, 929219 (2022).



25. Horowitz, K., Zayhowski, K., Palmour, N., Haghighat, D. & Joly, Y. Navigating the disclosure landscape: Parents' perspectives on healthcare professionals' role in supporting intersex children and families. *Journal of Genetic Counseling* doi:10.1002/jgc4.1962.
26. Statistics Canada. Income Explorer, 2021 Census. *Data Visualization, 2021 Census* <https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/dv-vd/income-revenu/index-en.html> (2022).
27. Statistics Canada Geography Division. Postal Code Conversion File [Canada], March 2024, Census of Canada 2021: Postal Code Conversion File, March 2024. Borealis <https://doi.org/10.5683/SP3/MR8S4U> (2024).
28. Statistics Canada. Statistical Area Classification (SAC). *Dictionary, Census of Population, 2021* <https://www12.statcan.gc.ca/census-recensement/2021/ref/dict/az/definition-eng.cfm?ID=geo045> (2021).