BMJ Open  JBI systematic review protocol of text/opinions on how to best collect race-based data in healthcare contexts

Cindy Quan,1 Nancy Clark,2 Catherine L Costigan,1 Jill Murphy,3 Michael Li,4 Anita David,5 Soma Ganesan,3 Jaswant Guzder,6 Barbara Cross4

ABSTRACT

Introduction  Racialized population groups have worse health outcomes across the world compared with non-racialized populations. Evidence suggests that collecting race-based data should be done to mitigate racism as a barrier to health equity, and to amplify community voices, promote transparency, accountability, and shared governance of data. However, limited evidence exists on the best ways to collect race-based data in healthcare contexts. This systematic review aims to synthesize opinions and texts on the best practices for collecting race-based data in healthcare contexts.

Methods and Analyses  We will use the Joanna Briggs Institute (JBI) methodology for synthesizing text and opinions. JBI is a global leader in evidence-based healthcare and provides guidelines for systematic reviews. The search strategy will locate both published and unpublished papers in English in CINAHL, Medline, PsycINFO, Scopus and Web of Science from 1 January 2013 to 1 January 2023, as well as unpublished studies and grey literature of relevant government and research websites using Google and ProQuest Dissertations and Theses. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement methodology for systematic reviews of text and opinion will be applied, including screening and appraisal of the evidence by two independent reviewers and data extraction using JBI’s Narrative, Opinion, Text, Assessment, Review Instrument. This JBI systematic review of opinion and text will address gaps in knowledge about the best ways to collect race-based data in healthcare. Improvements in race-based data collection, may be related to structural policies that address racism in healthcare. Community participation may also be used to increase knowledge about collecting race-based data.

Ethics and dissemination  The systematic review does not involve human subjects. Findings will be disseminated through a peer-reviewed publication in JBI evidence synthesis, conferences and media.

PROSPERO registration number  CRD4202368270.

INTRODUCTION

Health inequities for racialized groups have been a growing concern during COVID-19.1 The collection of ‘race’ data is controversial due to historical harms and ethical challenges.2 For example, projects such as the Human Genome Diversity Project, which attempted to map diverse populations, might perpetuate racism and the misbelief that ‘race’ captures biological differences among people.3 Advances in genetic research in the past decades show that ‘race’ is a social construct instead of having a genetic basis.4 Consistent with the scientific understanding of ‘race’ as a social construct, the field of human genetics has shifted to using alternative terms, such as ‘ancestry’ or ‘ethnicity’.5 6 In this review, we use the social constructs of ‘race’ as a social category to describe health disparities based on racism whereas ‘ethnicity’ describes cultural categories and grouping of people based on language, culture and/
or ancestry. We, therefore, refer to and define race-based data as the inclusion of sociodemographic data for healthcare as a process to mitigate racism and promote health equity.

The historical harm of treating ‘race’ as having a biological basis and the potential for using ‘race’ to discriminate has led to hesitations or misperceptions for collecting data about patients’ ‘race’ or ‘ethnicity’. For instance, healthcare providers perceive barriers to collecting demographic data, such as privacy and legal concerns, possible resistance from patients and staff, and difficulty recording the data. These attitudinal and practical barriers have led to the lack of clinic guidelines and less data collection. There is a need to correct misperceptions about the legality of data collection and bring clarity to best practices for collecting race-based data alongside other sociodemographic data.

Collecting race-based data is important to improve healthcare for all communities, especially for underserved, marginalized, and racialized communities. Racism is a social determinant of health and health inequalities. Health equity researchers agree that collecting disaggregated data in healthcare is in the best interest of diverse ‘racial’ or ‘ethnic’ groups. Thoughtful data collection conducted with an antiracism agenda can reduce health inequities for racialized communities.

Collecting data from racialized individuals requires competencies, such as attention to human rights, antiracism and culturally safe collection methods. For example, institutes for health research, such as the Canadian Institutes of Health Research, have addressed human rights issues for inclusion of sex and gender training to improve health disparities and transforming health outcomes for everybody. Structural policies such as these could be advanced by also including an Institute of Racism and Health.

In healthcare settings, there is important consideration for how information is collected, used, or disclosed for the purposes of identifying and eliminating systemic racism or to enhance individual care planning. Despite the importance of collecting race-based data, it is infrequently done or often poorly done when attempted. In a recent systematic review, Pan et al assessed reports of ethnicity and clinical outcomes of patients with COVID-19. Of the studies that included race or ethnicity data, Pan et al found an increased risk of acquiring SARS-CoV-2 infection among black, Asian and Minority Ethnic compared with white individuals and worse clinical outcomes from COVID-19. Their search included articles in English from studies around the world; only 12 of 690 articles identified from medical journals and 34 of the 209 preprints reported the ethnicity of patients. Since COVID-19, healthcare organisations are recognizing the need to improve the collection of disaggregated data to inform better healthcare services and support for patients. For example, before COVID-19, no Canadian province or territory systematically collected ‘race’ or ‘ethnicity’ data to identify and address the health disparities. Yet, the disproportionate risk and burden of COVID-19 among racialized communities has spurred a call from Canadian health and social care organizations to develop a nationwide strategy on collecting race-based and sociodemographic data. Moreover, race-based data collection in the USA and UK has shown that race-based data can be used to identify racial disparities and invest in interventions to address them, such as extended hours, new service locations, or extra staff. In countries that collect ‘race-based’ data, such as the USA, there are many problems with missing or inaccurate data. According to the Centers for Disease Control and Prevention, race and ethnicity data were missing for almost 40% of people testing positive for COVID or receiving a vaccine. Systematic reviews that have examined the reporting of race and ethnicity variables in medical research have found suboptimal reporting, including under-reporting of methods for establishing racial/ethnic categories, which leads to difficulties in interpreting and applying findings to support better health outcomes. Reporting on race data should include an analysis of the intersecting dimensions so that services can address the socially constructed categories which shape inequality and advance social justice. An intersectional approach allows for analyzing how complex social contexts impact the health and well-being of marginalized groups and pathways to care.

In the Canadian context, a new legal framework called the Anti-Racism Data Act (Bill 24-2022) became law in 2022. This framework provides a starting point for considering how information is collected, used, or disclosed for the purposes of identifying and eliminating systemic racism and advancing racial equity. The Anti-Racism Data Act has also spurred a call for action for all jurisdictions across Canada to start collecting ‘race-based’ data to address the disproportionate impacts of COVID-19 on racialized groups. The province of British Columbia (BC), for example, has recently adopted The BC Office of the Human Rights Commissioner (BCOHRC; 2020) Disaggregated demographic data collection: The Grandmother Perspective. Rooted in Indigenous values of data governance and social justice, this policy recommends that the methods for ‘race-based’ data collection must meaningfully include people directly impacted by this policy. Similar to the BCOHRC, Ontario Health, the governing body of Ontario’s healthcare system, recommends that public health systems standardize the collection of sociodemographic data to identify and address systemic causes of health inequities. Similarly, the National Institute on Minority Health and Health Disparities is 1 of 27 institutes and centres of the National Institutes of Health (NIH) that collaborates with communities and showcases the social determinants of minority health and health disparities within the US Department of Health and Human Services. Ultimately, initiatives to strengthen race-based data collection must include community-based participation and engagement in the collection and governance of race-based data.

Although some policies have provided some guidance and recommendations towards collecting race-based
data, little evidence exists on uptake of these practices and how it contributes towards improving equity. Policies and recommendations, therefore, do not always translate to high uptake or good data collection.31

There is a need for understanding if there is international consensus on the best approaches in collecting race-based data. Data collection is defined as the process of systematically gathering and measuring information on variables of interest to answer research questions, test hypotheses, and evaluate outcomes.37 Few rigorous primary research studies examine how best to collect race and ethnicity data in healthcare settings making quantitative systematic review inappropriate. Despite the trends to start collecting race-based data systematically within Canada and internationally, limited evidence exists on how this is being taken up. A systematic review of the best available text and opinion-based evidence from experts is well suited for supporting race-based data collection for healthcare, eliminating systemic racism, and advancing racial equity.34–36

Expert opinion derived by practitioners who typically are trained in the scientific method is often the best evidence in the absence of empirical studies. Joanna Briggs Institute (JBI) reviews about text/opinion do not draw on published research as the principal design of interest, rather they draw from areas of clinical care and tacit knowledge and discourses.38–40 We will use the JBI methodology for synthesizing narrative texts about policy, practice and opinions on collecting race-based data to provide a better understanding about the discourse on race-based data and provide guidance to researchers, practitioners and policy-makers. The JBI methodology for systematic review of text and opinion narratives evaluates diverse forms of knowledge including clinical tacit knowledge, cultural values, health policy and ethics.39 This includes grey literature sources such as, consensus reports, current discourse, comments, assumptions or assertions that appear in various journal commentaries, monographs, relevant websites and reports.40

The objective of the systematic review will be to provide recommendations, guidance, and insights on how to best collect race-based data in healthcare. The overarching research question is, ‘What are the best ways to collect race-based data within a healthcare context to improve services and equitable care?’

METHODS AND ANALYSIS

This protocol was developed following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA) guidelines (see online supplemental file 2).41 The systematic review was registered with PROSPERO. Reporting of the systematic review will be informed by PRISMA guidance.42 Amendments to the protocol will be documented and published with the results of the systematic review. The protocol will be conducted in accordance with the JBI methodology for systematic reviews of text and opinion. We will use the population, phenomena of interest, context, outcome, and time frame (PICOT) framework to guide the search strategy. This search will include grey literature sources. A preliminary search of PROSPERO, MEDLINE, the Cochrane Database of Systematic Reviews and the JBI Evidence Synthesis was conducted. No current or underway systematic reviews on the topic were identified.

Population

In this context, the population (P) does not imply aspects pertinent to quantitative reviews such as sampling method or sample size or homogeneity.40 The review will include text and opinions published by clinicians, healthcare organizations, administrators and policy-makers about their opinions on how to collect race/ethnicity data in a healthcare setting. We use health ‘care organisations’ to refer to healthcare organizations including hospitals, community health centres, administrative healthcare systems, medical home networks, public health agencies, etc. We will include published papers in English from American, British, Australian, and Canadian settings.

Intervention

Intervention (I) for our review includes broad areas of practice management to specific interventions related to race-based data collection such as administrative protocols within healthcare organizations. It is also noted that JBI text/opinion reviews may reflect on policies related to other aspects of healthcare, such as issues around equity and racism. In the intervention category, we are including opinions about health systems/services and delivery structures that promote healthcare equity.

Inclusion and exclusion criteria

Inclusion criteria of narrative texts for peer-reviewed and grey literature are: (1) standards for collecting race/ethnicity data; (2) consensus guidelines; (3) expert consensus; (4) narrative case reports; (5) published discussion papers; (6) conference proceedings and (7) government policy reports accessed from webpages of professional organizations.

Exclusion criteria: Technical reports, statistical reports, and epidemiological reports will be excluded. Opinion and media articles presenting opinions written by authors with no experience collecting or using race data, such as journalists, will be excluded.

Comparison/context

For JBI text/opinion systematic reviews, the use of a comparator is not required. However, the context in which the narrative text occurs will be included as it is relevant for our research question. In this review, we will include narrative texts that provide recommendations for how to collect race and ethnicity data to advance health equity in both public and private healthcare settings.

Outcome

A specific outcome is not required for JBI opinion/text reviews. But we will include a synthesis of best practices.
for collecting race-based data in healthcare contexts, such as clinical or administrative across specific settings, for example, hospital, or community. We adopt the recommended format by MacArthur et al for primary and secondary outcomes. The primary outcome is a synthesis of opinion about how race data are best collected to improve services and equitable care. The secondary outcome of interest in this review is a synthesis of reported (A) changes to the health system, (B) changes in cultural practice and (C) impact on health.

**Time**

Time describes the duration of our data collection. Because COVID-19 exposed growing health disparities for racialized population groups and the increased opinions about race-based data collection in this context, we will include narrative texts pre-COVID-19 and during COVID-19 (2013–2023) a 10-year time frame. Grey literature is also important to text/opinion reviews and will include relevant government research websites and contacting relevant organizations.

**Search strategy**

The search strategy aims to locate both published and unpublished papers. A three-step search strategy will be used in this review.

First, an initial limited search CINAHL (EBSCO) will be undertaken to identify articles on the topic. The text words contained in the titles and abstracts of relevant articles, and the index terms used to describe the articles will be used to develop a full search strategy for CINAHL, Medline, and PsycINFO, Scopus and Web of Science. Below is the search term used for CINAHL with restrictions set to English language, 1 January 2013 to 1 January 2023. See online supplemental appendix A for CINAHL search results and Scopus search.

**CINAHL Search:** collect* N5 (((ethnic* OR race OR racial* OR “race-based” OR “ethnic*-based” OR “ethnic* groups”) AND “race-based data collection” OR “Social Determinants of Health” OR (MH “Data Collection Methods”) OR (MH “Data Collection”)) OR “demographic data” OR “Personal Health Data” OR data AND health care OR “health care” OR hospital* OR “health service*” OR MH “Health Facilities*” OR health*.

The search strategy, including all identified keywords and index terms, will be adapted for each included database and/or information source. The reference list of studies included in the systematic review will be screened for additional studies. Studies published in English will be included. We will include studies published between January 2013 and January 2023. Second, we will search for sources of unpublished studies and grey literature using Google and ProQuest Dissertations and Theses database. For the Google search, we will use different keywords (eg, race-based and data and health and collecting; Personal Health Data and collecting; Social Determinants of Health; collecting and ‘demographic data’ and health) and screen the first 50 or until the relevance drops. As the ProQuest Dissertations and Theses database is set up like the peer-reviewed databases, we will use a similar initial search term as described in the keyword search above.

Third, we will also conduct targeted search of relevant national and international government and research websites, and for specific guidelines/consensus statements, on relevant society webpages.

**Government and research websites:**
- WHO.
- British Columbia’s Office of the Human Rights Commissioner.
- Wellesley Institute, Robert Wood Johnson Foundation.
- The Commonwealth Fund.
- National Institutes of Health.
- UK Government’s standards for ethnicity data.
- The Health Equity Collaborative Network.
- National Collaborating Centre for Determinants of Health.
- Black Dog Institute.

**Relevant society webpages:**
- Canadian Institute on Health Information’s section on race-based and Indigenous data identity.
- American Medical Association.
- American Psychological Association.
- Canadian Medical Association and Public Health Guidelines.

**Study selection criteria**

Following the search, all identified citations will be collated and uploaded into Covidence, a screening and data extraction tool, for screening and removal of duplicates. Titles and abstracts will then be screened by two or more independent reviewers for assessment against the inclusion criteria for the review. Potentially relevant papers will be retrieved in full, and their citation details imported into the JBI System for the Unified Management, Assessment and Review of Information (JBI-SUMARI). The full text of selected citations will be assessed in detail against the inclusion criteria by two or more independent reviewers. Reasons for exclusion of papers at full text that do not meet the inclusion criteria will be recorded and reported in the systematic review. Screening of titles, abstracts and full texts will be completed by two independent reviewers. Any disagreements that arise between the reviewers at each stage of the selection process will be resolved through discussion, or with an additional reviewer/s. The results of the search and the study inclusion process will be reported in full in the final systematic review and presented in a PRISMA flow diagram.

The nature of textual or opinion-based reviews is that they do not rely on evidence in the form of primary research and, therefore, elements of the protocol will aim to address a gap in primary reviews and highlight national and international policy and practice initiatives in relation to discourses on race-based data. Some text/opinion reviews that have included both primary studies...
Data will be extracted from studies included in the review by two independent reviewers using the standardized data extraction tool available in JBI-SUMARI. The data extracted will include specific details about the population, phenomena of interest and context relevant to the review objective. Conclusions will be extracted verbatim and assigned a level of credibility. To minimize errors, reviewers will read and reread each paper closely to identify the conclusions to be generated into JBI-NOTARI critical appraisal checklist. Any disagreements that arise between the reviewers will be resolved through discussion, or with a third reviewer.

**Assessment of methodological quality**

Eligible papers will be critically appraised by two independent reviewers for methodological quality using the standard JBI Critical Appraisal Checklist for Text and Opinion (JBI-Narrative, Opinion, Text, Assessment, Review Instrument (NOTARI)) developed by the JBI. The two reviewers will meet and discuss the criteria to ensure a common understanding and then apply the JBI-NOTARI to each piece of textual data until we reach a mutual decision about the quality of the text. See online supplemental appendix B for data extraction instrument for checklist used.

To determine whether an author of a paper has ‘experience in data collection’ or ‘have a standing in the field’ will be established by examining the credentials of the author. For instance, through their educational attainment (eg, degree in a relevant field), vocation (eg, mental health clinician, health informatics specialist) or lived experience (eg, advocate in a mental health organisation). Expert opinions are often mediated through practitioners trained in scientific methods. However, in our systematic review, we also consider the authenticity of the opinion and alternative opinions. Specifically, we will also consider equity, diversity, and inclusion (EDI) contexts or training and application of other EDI frameworks, for example, intersectionality.

To establish credibility of the papers under consideration for inclusion, we will also use three levels of credibility to help determine whether or not conclusions represent the evidence, this includes (1) Unequivocal (U): Relates to evidence beyond reasonable doubt which may include conclusions that are matter of fact, directly reported/observed and not open to challenge; (2) Credible (C) relates to those conclusions that are interpretative, or plausible in light of the data or theoretical framework and (3) Unsupported (U) when findings are not supported by the data.

Authors of papers will be contacted to request missing or additional data for clarification, where required. Any disagreements that arise will be resolved through discussion, or with a third reviewer. The results of critical appraisal will be reported in narrative form and in a table. All papers, regardless of the results of their methodological quality, will undergo data extraction and synthesis. The results of the critical appraisal of papers will be reported in a table and in narrative form.

**Data synthesis**

Textual papers will, where possible, be pooled using JBI-SUMARI with the meta-aggregation approach. This will involve the aggregation or synthesis of conclusions to generate a set of statements that represent that aggregation, through categorising assembled conclusions based on similarity in meaning. These categories will then be subjected to a synthesis in order to produce a single comprehensive set of synthesised findings that can be used as a basis for evidence-based practice. Where textual pooling is not possible, the findings will be presented in narrative form. Only unequivocal and credible findings based on the JBI-NOTARI critical appraisal checklist will be included in the synthesis.

**Assessing confidence in the findings**

The final synthesized findings will be graded according to the ConQual approach for establishing confidence in the output of textual synthesis and presented in a Summary of Findings table. The Summary of Findings table includes the major elements of the review and details how the ConQual score is developed. Included in the table is the title, population, phenomena of interest and context for the specific review. Each synthesized finding from the review is then presented along with the type of research informing it, a score for dependability, credibility and the overall ConQual score.

**ETHICS AND DISSEMINATION**

The systematic review does not involve human subjects. Findings will be disseminated through a peer-reviewed publication in JBI evidence synthesis, conferences and media.

**Patient and public involvement**

Patients and members of the public are involved at all stages of the research. One coauthor is a patient partner with expertise in community-based research approaches. They contributed to the design of the study and will be involved with the analyses and dissemination of the study findings. Several coauthors are public partners working in a large health organisation network. The partners working in the healthcare system contributed to the design of the review and will be involved in the analyses and development of clinical and practice guidelines based on the findings. As the study is a systematic review, there are no research participants.
DISCUSSION

Evidence and healthcare approaches play a major role in clinical decision-making. The synthesis of expert opinion reviews is not well recognized in mainstream evidence-based practice. In the absence of best available evidence about race-based data collection, using opinions in the form of texts or narratives can provide valuable recommendations for clinical practice. We concur with researchers who argue that we need to broaden the evidence to promote healthcare equity for all and particularly for racialized patients. Our review will add an important contribution to assisting policy makers, researchers, and clinicians to collect race-based data in practice. We note that there are historically many negative consequences related to race-based data collection. However, in the context of increasing health disparities among racialized population groups pre-COVID-19 and during COVID-19 there is an increased call for collecting disaggregated data that shapes health outcomes including health system responses to deliver equitable and quality care, The nature of textual or opinion-based reviews is that they provide a credible source of evidence in the absence of primary data. Race-based data have not been systematically collected across healthcare contexts and little is known about how this data ought to be collected. Our systematic review of text and opinions on how to best collect race-based data in healthcare contexts will produce a synthesis of conclusions aimed to guide healthcare practices. Given the growing health disparities between and within racialized population groups, we anticipate that the synthesized conclusions may have applicability across healthcare contexts and settings. Researchers have argued that health equity cannot be achieved without equitable evidence. Our synthesis will also add value to providing equitable evidence for best practice approaches to collecting race-based data in healthcare contexts.

Supplemental material

This content has been supplied by the author(s), It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

REFERENCES


Twitter Nancy Clark @clarkeucinurs

Acknowledgements We acknowledge the help and support of our Librarian Jessica Russell, University of Victoria and funding support provided by the Uvic Internal Creative Project Research Grant.

Contributors This systematic review was conceived and designed by all authors. NC, JM, CLC, JG and SG provided methodological expertise. AD, ML and BC provided expertise from lived experience or community perspective. CQ and NC drafted the protocol. Subsequent drafts were commented and NC developed the search strategy, as well as eligibility and data extraction and NC developed the search strategy, as well as eligibility and data extraction and BC provided expertise from lived experience or community perspective. CQ and NC drafted the protocol. Subsequent drafts were commented

Funding The authors declare mentorship and funding support from the University of Victoria (UVic) Centre for Evidence-Informed Nursing and Healthcare (CEINHC); A JBI Centre of Excellence CEINHC and the University of Victoria’s Internal Research Grant.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Ethics and dissemination section for further details.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.