

Selecting contextually appropriate performance indicators for Northern maternity care systems: A modified Delphi approach

Rebecca Rich (✉ rich@ualberta.ca)

University of Alberta <https://orcid.org/0000-0002-2048-2671>

Thomsen D'Hont

University of Alberta

Kellie Murphy

University of Toronto

Jeremy Veillard

The World Bank

Susan Chatwood

Institute for Circumpolar Health Research

Research article

Keywords: maternal health, infant health, Arctic regions, Indigenous health services, health care quality indicators

DOI: <https://doi.org/10.21203/rs.3.rs-49446/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Meaningful performance measurement requires indicators to be scientifically robust and strategically focused. For many circumpolar states, indicators aligned with national strategies may ignore the priorities of northern, remote, or Indigenous populations. The aim of this project was to identify contextually appropriate performance indicators for maternity care in Northern regions.

Methods

A scoping review generated a working list of indicators. Fourteen maternity care and health systems experts then participated in a modified Delphi consensus process. Participants rated proposed indicators according to importance, circumpolar relevance, validity, and reliability and suggested additional indicators for consideration.

Results

Consensus was achieved after two rounds. Eleven indicators were rated highly on all four criteria. Twenty-nine indicators, largely focused on social determinants of health, health care responsiveness, and accessibility, were identified for further research.

Conclusions

The Delphi approach was effective in identifying indicators for maternity care in Northern regions. Further research and collaboration with key stakeholders is necessary to further develop indicators of health system responsiveness and accessibility.

Background

Among health systems around the world, performance measurement has become an increasingly popular tool in the pursuit of health care quality, accountability and value for money (1–4). This trend has been accompanied by an explosion in the number of available performance indicators and a dramatic increase in the number of organizations collecting and reporting on health system performance information. In this environment, meaningful performance measurement requires indicators that are scientifically robust and strategically focused. Indicators must be valid, reliable, sensitive to change, and feasible to measure. They must also be aligned with the values and priorities of the system they are intended to evaluate (5–7).

For circumpolar states, those whose political borders encompass the most northern regions of the globe, the selection of contextually relevant indicators presents unique challenges. Geopolitically, “circumpolar” refers to all of Alaska, Greenland, Iceland, the Faroe Islands, and the northernmost territories of Sweden, Finland, Norway, Russia and Canada. These regions are characterized by vast distances, harsh winter climates, and low population densities, all of which make the delivery of health care challenging and expensive. These health systems thus rely more heavily on long distance travel of both patients and providers than do their southern counterparts. In addition, many circumpolar regions have large Indigenous populations which may comprise up to 90% of the population. It is thus imperative to consider the historical and ongoing impacts of settler colonialism and the substantial health inequities that exist both within and between circumpolar regions and disproportionately affect Indigenous peoples.

Maternity care provides a useful example of the unique health care landscape in Northern regions. While medical evacuation can provide critical access to secondary and tertiary level care and is an important component of all rural and remote health systems, efforts at centralization have also put the sustainability of local birth programs at risk. The resultant practice of routinely evacuating low-risk women for labour and birth has substantial and well-documented consequences for women, families and communities (8–11). Where possible, it is considered best practice to provide care for labour and birth closer to home (12, 13). Being a uniquely rural/remote phenomenon, this issue affects a minority of women in Canada and other circumpolar states and is not often captured in pan-national performance measurement efforts (14, 15). Canadian maternity care indicators, for example, focus largely on decreasing practice variation with respect to obstetrical interventions, such as induction of labour, operative vaginal delivery and caesarean section. This approach has positioned cost savings and improvements in patient safety through a reduction in intervention as central performance targets (16). Pan-national indicators such as these are often selected based on the strategic priorities of a dominant southern majority and may ignore or even conflict with the needs and priorities of northern or remote communities. This is particularly relevant for health systems serving Indigenous communities as the importance of including Indigenous knowledge and stakeholder input in the selection of performance indicators has been well documented in the literature (17, 18). This study attempts to begin filling this gap by identifying performance indicators which are contextually appropriate for Northern health systems.

Methods

A two-round modified Delphi consensus study was conducted among Northern maternity care experts. The objectives of the study were to establish consensus on indicator priorities and identify contextually relevant perinatal performance indicators. The Delphi method is frequently applied in circumstances where the available evidence is contradictory or insufficient (19) and is a common method for the selection of key performance indicators (1, 20-23).

This study was carried out in accordance with the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans. Particular attention was devoted to chapter 9, Research Involving the First

Nations, Inuit and Métis Peoples of Canada (24). There was consultation throughout the project with a First Nations Elder, medicine woman and healer who has extensive experience working with health researchers. This study was approved by the University of Toronto research ethics board. Territorial research licenses were obtained from the Aurora and Nunavut Research Institutes.

Participant selection

Participants were selected using purposeful sampling as the quality of results produced by a Delphi panel depends on the expertise of participants (25). We sought out a heterogeneous panel to ensure that a wide variety of stakeholder groups were represented. A target sample size of 10-20 was chosen to balance the trade-offs between decision quality and data manageability as well as between panel heterogeneity and the ability to achieve consensus.

Derivation of survey items

The questionnaires used was developed for this consensus process. Survey items were derived from an extensive literature review, which has been published elsewhere (15). Sixty-two unique maternity care performance indicators were identified through this review and subsequently subdivided into the following domains according to a modified version of the OECD Health Performance Framework (26): Determinants of health, health outcomes, health system effectiveness, safety, responsiveness, accessibility, and cost.

Response scale and criteria

Respondents were asked to rate each of the 62 indicators on a 7-point Likert-type scale ranging from 1 (strongly disagree) to 7 (strongly agree) according to the following four criteria:

- Importance: The level of concern of health care users or policy makers and the degree to which the indicator is susceptible to influence by the health system.
- Circumpolar relevance: The significance of the indicator within the participant's circumpolar context.
- Validity: The degree to which the indicator measures what it intends to measure.
- Reliability: The degree to which the indicator provides stable results across various populations, circumstances, and time points.

In order to allow participants to best assess the validity and reliability of each indicator, a detailed information package was provided to each participant. During the first round, participants were also asked to suggest additional indicators not identified in the literature review. These additional indicators were incorporated into the second-round survey for evaluation by the panel[1].

Survey administration

The questionnaires were created and distributed electronically. Two researchers piloted the questionnaire to ensure it was clearly worded and the suggested time for completion was appropriate. Responses to the first-round survey were collected over a period of 41 days. Responses were collected and tabulated to create the second-round survey, which included the same 62 indicators along with the 17 additional indicators suggested by participants. Round-two responses were collected over a period of 19 days.

Data analysis

After collection of the first-round responses, measures of central tendency and dispersion (mean, median and standard deviation (SD)) were calculated for each indicator against each criterion. This information was included as structured feedback in the second round of the survey. Consensus among participants was determined by the degree of internal consistency demonstrated by participant responses. This was established using standardized Cronbach's alpha. This method assumes that each indicator possesses a true level of each characteristic. Each participant's response to a given survey question represents a single measurement of that indicator's importance, relevance, validity, or reliability. The internal consistency of responses can thus be considered a measure of agreement among participants. Published thresholds for an appropriate Cronbach's alpha vary in the literature from 0.70 (27) for the determining the inter-rater reliability of a psychometric scale designed to differentiate between groups to 0.90 (28) for use of a diagnostic scale on an individual level. The threshold for determining consensus among experts on a Delphi panel is similarly variable but a cut-off of a ³ 0.8 is frequently used (29, 30). All statistical calculations were performed using SAS University Edition.

Selection of key indicators

After achievement of consensus, indicators for which ³80% of participants selected 6 (agree) or 7 (strongly agree) on all four criteria were identified as a core set of indicators. Using importance and circumpolar relevance as gateway criteria, a second set of indicators was also identified for consideration. The latter group included indicators for which ³80% of participants provided a strong rating (6 or 7) for the criteria of importance and relevance but that did not meet the same threshold for validity and reliability.

[1] The second round survey has been included as an additional file.

Results

Achieving consensus

Twenty-one experts were invited to participate in the Delphi consensus process. Fourteen experts (67%) completed two rounds of the Delphi process. The group included two representatives from Alaska, one from the Yukon Territory, two from the Northwest Territories, four from Nunavut, two from Greenland, and

three who currently work in southern Canada but have significant expertise in Arctic or circumpolar maternal and child health. The panel included seven physicians (two obstetricians, three family physicians, one pediatrician, and one medical geneticist), four midwives, one nurse, and two public health researchers. Many of the participating clinicians currently work in research or health policy roles, allowing them to bring multiple areas of expertise to the panel.

The level of agreement among participants, as measured by standardized Cronbach’s alpha, is displayed in Table 1. After two rounds, the level of agreement was very good as demonstrated by a Cronbach’s alpha of 0.87. Missing values accounted for 1.3% of all response options.

Table 1. Level of Agreement Among Delphi Participants

Round	Participants	Agreement (Cronbach’s a)		
		Missing values excluded	Missing values replaced with neutral value (4)	Missing values replaced with mean value
Round 1	14	0.79	0.75	0.79
Round 2	14	0.87	0.86	0.87

Standardized Cronbach’s a calculated based on management of missing data. Consensus achieved when $a > 0.8$.

Selected indicators

Once consensus had been reached, the proportion of responses indicating participant agreement (6 - agree or 7 - strongly agree) was tabulated for each survey statement. Eleven indicators were rated highly on all four criteria by ³ 80% of the participants and have been labeled the “core” indicators (Table 2). This set of indicators primarily focuses on physical health outcomes but does contain four indicators which represent other domains of health system performance. These include teenage pregnancy and birth, admissions to the neonatal intensive care unit (NICU), presence of a skilled birth attendant, and travel to place of birth.

Table 2. Selected “Core” Performance Indicators

Indicator	Ratings ³ 6 (%)			
	Importance	Relevance	Validity	Reliability
<u>Determinants of health</u>				
Teenage pregnancy & birth	93	93	93	93
<u>Health outcomes</u>				
Anemia in pregnancy	93	93	86	86
Stillbirths	100	93	93	93
Perinatal deaths	100	100	100	100
Preterm birth	100	100	86	86
Maternal Mortality	100	93	100	100
Neonatal mortality	100	100	100	100
Low birth weight	86	93	93	93
<u>Effectiveness</u>				
NICU admissions	100	85	100	100
<u>Accessibility</u>				
Birth attendant	93	93	93	93
Travel to place of birth	92	93	86	86

Indicators for which >80% of participants provided a rating of 6 or 7 for each criterion.

Twenty-nine additional indicators received high ratings for importance and relevance but lower ratings for reliability or validity, indicating that further study or development of these indicators may be required (Table 3). It is important to note that this group of indicators includes a much greater focus on social determinants of health, health system responsiveness and health care accessibility. Three major themes are represented within this second group of indications: Increased focus on the social determinants of health; the impact of medical travel; and the importance of culturally appropriate care at both provider and system levels.

Table 3. Selected “Additional” Performance Indicators

Indicator	Ratings ³ 6 (%)			
	Importance	Relevance	Validity	Reliability
<u>Determinants of health</u>				
Maternal education level	93	93	93	79
Domestic violence	93	100	21	7
Smoking during pregnancy	100	100	57	50
Use of illicit drugs during pregnancy	100	93	21	14
Use of alcohol during pregnancy	100	100	29	21
Breastfeeding practices	100	100	64	43
Food insecurity*	100	100	-	-
Maternal housing (crowded or under-housed)*	93	93	-	-
<u>Health outcomes</u>				
Diabetes in pregnancy	93	93	86	79
Postpartum hemorrhage	93	93	71	50
Postpartum depression	86	93	29	21
Congenital anomalies	86	86	86	79
Small for GA infants	100	100	86	79
<u>Effectiveness</u>				
Neonatal readmission to hospital	86	86	93	79
<u>Safety</u>				
Births without obstetric intervention	93	86	64	64
Transfers for obstetrical indications*	92	85	-	-
Unplanned births in the community*	100	92	-	-
<u>Responsiveness</u>				
Characteristics of care providers	92	93	69	69
Cultural competency	92	86	46	38
Patient reported unfair treatment based on ethnicity	92	86	57	29
Patient reported support during labour and birth	100	86	50	50

Indicator	Ratings ³ 6 (%)			
	Importance	Relevance	Validity	Reliability
Presence of breastfeeding support programs	92	86	29	14
Patient reported satisfaction with care	100	100	50	36
<u>Accessibility</u>				
Frequency and timing of antenatal care	93	93	79	71
Use of antenatal ultrasound	86	86	93	79
Rate of induced abortions	86	86	86	79
Postpartum visits	100	86	71	64
Maternity care provider in patient's community*	100	93	-	-
Maternity care provider who speaks the same language/is from the same culture as the patient*	92	93	-	-

Indicators for which $\geq 80\%$ of participants provided a rating of 6 or 7 on the basis of importance and circumpolar relevance.

**Denotes additional indicator suggested by members of the Delphi panel in the first round.*

Discussion

This study demonstrates the ability of a modified Delphi approach to achieve consensus among a geographically dispersed and professionally diverse group of stakeholders. Through this structured consensus process, we selected a core group of performance indicators and identified other important indicators for further consideration and development. The eleven core indicators (Table 2), are largely focused on morbidity and mortality. This is consistent with the findings of our earlier work (15) and reflects the historical dominance of physical health outcomes in performance measurement and health surveillance systems. Physical health outcomes provide the most straightforward opportunity for standardizing definitions and collecting high quality data. The validity and reliability of these indicators is also well established in the literature. A more interesting and impactful outcome of the study is the identification of 29 additional indicators for further consideration (Table 3). These indicators, which were identified using importance and relevance as gateway criteria, demonstrate interest in evaluating Northern health system performance through the domains of social determinants of health, healthcare accessibility, and health system responsiveness. Themes related to travel for care and the availability of culturally appropriate care weave throughout each domain.

Social and structural determinants of health such as income, education, social support, gender, and experiences of racism are all known to influence health behaviours, health system access and health

outcomes (31–34). Important social determinants of health highlighted in this study include maternal education, the presence or absence of intimate partner violence, food insecurity and inadequate housing. While there is strong evidence that Northern communities are disproportionately affected by these challenges, they are not routinely included in perinatal health surveillance.

In Alaska, the Northern Canadian Territories and Greenland, the structural, political, and social determinants of health are compounded by physical isolation, increased industrialization, changing environmental conditions, and long winters. These factors contribute to the geographical and logistical barriers that Northerners face when accessing healthcare. Multiple indicators of healthcare access were highlighted in this study. These include but are not limited to the number of antepartum or postpartum visits a patient is able to attend, the availability of obstetrical ultrasound, and the distance a patient is required to travel for care. The latter indicator was identified from the Canadian Maternity Experiences Survey (35) and captures the proportion of women who report traveling > 100 km to receive care. It is important to remember that in many Northern regions around the world, distances are almost always greater than 100 km. With this in mind, the Delphi panel suggested other measures of travel for care including the proportion of women who can access a maternity care provider (and/or skilled birth attendant) in their community, the number of unplanned births occurring in each community (which generally occur as a result of patients being unable or unwilling to travel for birth), and the number of emergency transfers for obstetrical indications. If “distance” to care is measured within a Northern context, we suggest using the number of “days or hours away from home” as a more appropriate reflection of the travel burden. The Delphi panel also suggested that an indicator be developed to capture the proportion of women in a region receiving care in their own language. This, of course, reflects the fact that accessibility is not equivalent to geographical proximity and equitable access to care requires that systems and providers minimize social, financial and systematic barriers to care and ensure the availability of culturally safe care for Indigenous women.

Health system responsiveness is a concept intended to capture the ability of a health system to acknowledge, accommodate, and react to the expectations of the population. It shares many tenants with patient centered care and cultural competency (36). For Indigenous women, providing culturally competent and thus responsive care may necessitate a holistic appreciation of health (37, 38) including physical, mental, emotional, spiritual, and historical considerations. It is well established that racism, colonization, and self-determination are important determinants of Indigenous peoples’ health and are major drivers of health inequities (10, 39, 40). The Delphi panel demonstrated this critical issue by selecting multiple indicators reflective of cultural competency or safety or other aspects of the patient experience. These indicators include a number of patient reported outcomes in order to reflect the importance of assessing cultural safety through the eyes of the individual, family or community receiving care. Our study suggests that Northern health care providers, institutions and systems should not only prioritize efforts to provide culturally safe care, but should also measure and report on their ability to do so.

While many of the selected indicators are not ready for implementation as they have been presented here, Northern health regions interested in improving their use of contextually relevant performance indicators, may find this a useful resource. This study was initially intended to take on a circumpolar approach. However, due to the geographical distribution of participating experts, it can more accurately be applied to Alaska, the Northern parts of Canada, and Greenland. Of course, other Northern or rural/remote regions may find that our findings can be adapted for their context. As presented, many of indicators outlined in this study do not currently demonstrate the scientific rigor to allow for immediate use. This is due to a variety of factors including the complex nature of the constructs being measured, the availability of data sources to support these indicators, and the willingness of health systems leadership to direct resources toward development of indicators for small populations. In order to prepare these indicators for use at local, regional or national levels, we recommend that further indicator development be carried out in consultation with patients, families, community leaders as well as other local stakeholders. Concomitant development and application of region-specific health performance frameworks will help to maintain comprehensiveness and relevance while minimizing indicator redundancy.

Beyond the specific indicators identified by the Delphi panel, our findings highlight some important challenges and opportunities for Northern performance measurement initiatives. First, when reporting rare outcomes among small populations, aggregation of data over large regions or timeframes may be required to ensure confidentiality and statistical rigor. Thus, there may be a larger role for process indicators and for confidential audits of rare outcomes such as maternal and neonatal mortality. Second, health system performance measurement is only as effective as the data that supports it. In the Canadian territories, for example, there is limited prospective collection of perinatal surveillance data. There is also a marked lack of Indigenous specific identifiers within Canadian health data sources more broadly (41, 42). The development and implementation of Indigenous health information systems in Canada is essential and must be carried out with and by Indigenous communities and organizations to ensure appropriate indicator selection, data usage and governance (43, 44). Finally, and most importantly, our findings reinforce the fundamental role of context and values within health system stewardship (6, 17, 45) and the importance of aligning health policy with broader agendas (46). In Canada, we have seen the recent adoption of the UN Declaration on the Rights of Indigenous Peoples and the release of the final reports of the Truth and Reconciliation Commission and the Inquiry into Missing and Murdered Indigenous Women and Girls (47–50). In association with these reports, Canadian National bodies have begun to recognize the importance of Indigenous rights and self-governance. Despite this, the commonly used performance measurement systems are still incongruent with these values. Canadian policy makers thus have an opportunity and an obligation to engage with northern stakeholders and Indigenous communities to ensure that health policies and resources are directed appropriately. Furthermore, when developing performance measurement strategies intended to capture and address health inequities, it is critical to recognize the role that defect-based approaches to health surveillance have had and continue to have on the stigmatization of Indigenous peoples within the health literature. Adequate engagement and attention to strengths-based approaches can limit this unintended consequence while simultaneously providing opportunities for health system improvement and management.

Conclusion

In this study, we identified eleven scientifically robust maternity care performance indicators that are important and relevant. We also identified many indicators that warrant further consideration and development as they more broadly reflect the priorities in Northern regions. Based on the findings of this study, we recommend that Northern regions re-evaluate their current perinatal performance measurement systems to ensure contextual appropriateness. It is our hope that development and use of contextually appropriate indicators will foster continuous quality improvement in Northern maternity centers, offer opportunities for inter-regional comparisons, and compel the development and improvement of Northern and Indigenous health information systems.

Abbreviations

OECD – Organization for Economic Cooperation and Development

NICU – neonatal intensive care unit

GA – gestational age

NWT (or NT) – Northwest Territories

NU – Nunavut

Declarations

Ethics

This study was carried out in accordance with the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans. Particular attention was devoted to chapter 9, Research Involving the First Nations, Inuit and Métis Peoples of Canada (24). There was consultation throughout the project with a First Nations Elder, medicine woman and healer who has extensive experience working with health researchers. This study was approved by the University of Toronto research ethics board. Territorial research licenses were obtained from the Aurora and Nunavut Research Institutes. Informed consent was obtained from each participant. Participants demonstrated their consent by initiating responses to the voluntary online survey.

Consent for Publication

Not applicable

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request

Competing interests

The authors declare that they have no competing interests

Funding

This work of author TD was supported by the Canadian Institutes of Health Research team grant in community-based primary health care (No. TT6-128271). Community engagement and knowledge translation was supported by the Canadian Foundation for Women's Health (Dawn Walker Grant) and the Canadian Institutes of Health Research – Strategy for Patient Oriented Research patient engagement – collaboration grant (No. 201511PEG). The funding agencies had no role in the design, conduct, or interpretation of the study.

Authors' contributions

RR designed the study, collected and analyzed the data, and drafted the manuscript. TD contributed to data collection, analysis, and manuscript editing. KM, JV, and SC assisted RR in all aspects of the project in their role on RR's graduate studies supervisory committee. All authors have read and approved the manuscript.

Acknowledgements

The authors wish to express their sincere gratitude to Elder Be'sha Blondin, as well as the NWT community members and anonymous reviewers of the NWT and NU research licensing bodies for their contextual contributions to this study. We would also like to extend our deep appreciation to the Delphi panel participants who generously provided their time and expertise.

References

1. Carinci F, Van Gool K, Mainz J, Veillard J, Pichora E, Januel J, et al. Towards actionable international comparisons of health system performance: expert revision of the OECD framework and quality indicators. *Int J Qual Health Care*. 2015;27(2):137–46.
2. Freeman T. Using performance indicators to improve health care quality in the public sector: a review of the literature. *Health Services Management Research*. 2002;15(2):126–37.
3. Murray C, Evans D. Health systems performance assessment: debates, methods and empiricism. Geneva: World Health Organization; 2003.
4. Smith PC, Mossialos E, Papanicolas I. Performance measurement for health system improvement: experiences, challenges and prospects. 2008.
5. Brown AD, Veillard J. Commentary. Indicators with a Purpose—Meaningful Performance Measurement and the Importance of Strategy. *Healthcare Policy*. 2008;3(4):49.
6. Veillard J, Huynh T, Ardal S, Kadandale S, Klazinga NS, Brown AD. Making health system performance measurement useful to policy makers: aligning strategies, measurement and local

- health system accountability in Ontario. *Healthcare Policy*. 2010;5(3):49.
7. World Health Organization. *The world health report 2000: health systems: improving performance*: World Health Organization; 2000.
 8. Chamberlain M, Barclay K. Psychosocial costs of transferring indigenous women from their community for birth. *Midwifery*. 2000;16(2):116–22.
 9. Lawford K, Giles AR. An analysis of the evacuation policy for pregnant First Nations women in Canada. *AlterNative: An International Journal of Indigenous Peoples*. 2012;8(3):329.
 10. Moffitt PM. Colonialization. A health determinant for pregnant Dogrib women. *J Transcult Nurs*. 2004;15(4):323–30.
 11. Varcoe C, Brown H, Calam B, Harvey T, Tallio M. Help bring back the celebration of life: A community-based participatory study of rural Aboriginal women's maternity experiences and outcomes. *BMC Pregnancy and Childbirth*. 2013;13.
 12. Lalonde A, Grzybowski SC, Klein MC, Iglesias S, Gagne GP. Rural obstetrics: joint position paper on rural maternity care. *Can J Rural Medicine*. 1998;3(2):75.
 13. Miller KJ, Couchie C, Ehman W, Graves L, Grzybowski S, Medves J, et al. Rural maternity care. *Journal of Obstetrics Gynaecology Canada*. 2012;34(10):984–91.
 14. Canadian Institute for Health Information (CIHI). *Canadian Institute for Health Information. Health Indicators 2013*. Ottawa: CIHI; 2013.
 15. Rich R, D'Hont T, Linton J, Murphy KE, Veillard J, Chatwood S. Performance Indicators for Maternity Care in a Circumpolar Context: A Scoping Review. *Int J Circumpolar Health*. 2016;75 (31470).
 16. Canadian Institute for Health Information (CIHI). *Canadian Institute for Health Information: Indicator Library 2015* [cited 2015 Nov]. Available from: <http://indicatorlibrary.cihi.ca/display/HSPIL/Indicator+Library?desktop=true>.
 17. Chatwood S, Bytautas J, Darychuk A, Bjerregaard P, Brown A, Cole D, et al. Approaching a collaborative research agenda for health systems performance in circumpolar regions. *Int J Circumpolar Health*. 2013;72.
 18. Anderson M, Smylie J, Anderson I, Sinclair R, Crengle S. *First Nations, Metis, and Inuit Health Indicators in Canada: A background paper for the project "Action Oriented Indicators of Health and Health Systems Development for Indigenous Peoples in Australia, Canada, and New Zealand"*. The Indigenous Peoples' Health Research Centre (IPHRC); 2006.
 19. Graham B. Delphi as a method to establish consensus for diagnostic criteria. *J Clin Epi*. 2003;56(12):1150–6.
 20. Zeitlin J, Wildman K, Breart G, Alexander S, Barros H, Blondel B, et al. PERISTAT: indicators for monitoring and evaluating perinatal health in Europe. *Eur J Public Health*. 2003;13(3 Suppl):29–37.
 21. Boulkedid R, Sibony O, Goffinet F, Fauconnier A, Branger B, Alberti C. Quality Indicators for Continuous Monitoring to Improve Maternal and Infant Health in Maternity Departments: A Modified Delphi Survey of an International Multidisciplinary Panel. *PLoS ONE*. 2013;8(4):1–11.

22. Gagliardi AR, Fung MFK, Langer B, Stern H, Brown AD. Development of ovarian cancer surgery quality indicators using a modified Delphi approach. *Gynecol Oncol.* 2005;97(2):446–56.
23. Boukchedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS ONE.* 2011;6(6):e20476.
24. Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada. Social Sciences and humanities Research Council of Canada. Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans. Ottawa, Canada; 2014.
25. Hsu C-C, Sandford BA. The Delphi Technique: Making Sense Of Consensus. *Practical Assessment, Research & Evaluation.* 2007;12(10).
26. Arah OA, Klazinga N, Delnoij D, Ten Asbroek A, Custers T. Conceptual frameworks for health systems performance: a quest for effectiveness, quality, and improvement. *Int J Qual Health Care.* 2003;15(5):377–98.
27. Nunnally JC, Bernstein I. The assessment of reliability. *Psychometric Theory.* 1994;3(1):248–92.
28. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. *BMJ.* 1997;314(7080):572.
29. Zevin B. Design and Validation of a Comprehensive Simulation-Enhanced Training Curriculum for a Complex Minimally Invasive Operation: University of Toronto; 2014.
30. Shore E. Design of a Comprehensive Evidence-Based Laparoscopy Curriculum for Gynecology Residents. Toronto: University of Toronto; 2014.
31. Canadian Institute for Health Information (CIHI). A performance measurement framework for the Canadian health system. Ottawa: CIHI; 2013.
32. Marmot M, Friel S, Bell R, Houweling TA, Taylor S, Commission on Social Determinants of Health. Closing the gap in a generation: health equity through action on the social determinants of health. *The Lancet.* 2008;372(9650):1661–9.
33. Dahlgren G, Whitehead M. Policies and strategies to promote social equity in health. Stockholm: Institute for future studies; 1991.
34. Paradies Y, Ben J, Denson N, Elias A, Priest N, Pieterse A, et al. Racism as a determinant of health: a systematic review and meta-analysis. *PloS one.* 2015;10(9):e0138511.
35. Public Health Agency of Canada. What Mothers Say: The Canadian Maternity Experiences Survey. Ottawa; 2009.
36. Saha S, Beach MC, Cooper LA. Patient centeredness, cultural competence and healthcare quality. *J Natl Med Assoc.* 2008;100(11):1275–85.
37. Smylie J, Anderson I, Ratima M, Anderson M. Indigenous health performance measurement systems in Canada, Australia, and New Zealand. *The Lancet.* 2006;367:2029–31.
38. Stephens C, Nettleton C, Porter J, Willis R, Clark S. Indigenous peoples' health—why are they behind everyone. everywhere? *The Lancet.* 2005;366(9479):10–3.
39. Allan B, Smylie J. First Peoples, second class treatment: The role of racism in the health and well-being of Indigenous peoples in Canada. Toronto: The Wellesley Institute; 2015.

40. Reading CL, Wien F. Health inequalities and the social determinants of Aboriginal peoples' health: National Collaborating Centre for Aboriginal Health Prince George, BC; 2009.
41. Walter M, Andersen C. Indigenous Statistics. A Quantitative Research Methodology. Walnut Creek: Left Coast Press, Inc.; 2013.
42. Smylie J, Firestone M. Back to the basics: Identifying and addressing underlying challenges in achieving high quality and relevant health statistics for Indigenous populations in Canada. *Statistical Journal of the IAOS*. 2015;31(1):67–87.
43. Mosby I. Administering colonial science: Nutrition research and human biomedical experimentation in Aboriginal communities and residential schools, 1942–1952. *Social history*. 2013;46(1):145–72.
44. Harry D, Kaneche LaM. Asserting tribal sovereignty over cultural property: moving towards protection of genetic material and indigenous knowledge. *Seattle J Soc Just*. 2006;5:27.
45. Saltman RB, Ferroussier-Davis O. The concept of stewardship in health policy. *Bull World Health Organ*. 2000;78(6):732–9.
46. Veillard JHM, Brown AD, Barış E, Permanand G, Klazinga NS. Health system stewardship of National Health Ministries in the WHO European region: concepts, functions and assessment framework. *Health Policy*. 2011;103(2):191–9.
47. Wolfe S. *The politics of reparations and apologies*: Springer Science & Business Media; 2013.
48. Truth and Reconciliation Commission of Canada. *Summary of the final report of the Truth and Reconciliation Commission of Canada*. Winnipeg: Truth and Reconciliation Commission of Canada; 2015.
49. UN General Assembly. *United Nations Declaration on the Rights of Indigenous Peoples*. 2007.
50. Girls NiiMaMIWa. *Reclaiming Power and Place: The Final Report of the National Inquiry Into Missing and Murdered Indigenous Women and Girls: National Inquiry into Missing and Murdered Indigenous Women and Girls*; 2019.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [DelphiQuestionnaireRound2.pdf](#)