# Structural Determinants of Indigenous Health Disparities in Canada: An Integrated MFA, SEM, and Machine Learning Analysis

Timothy Leung

October 22, 2025

## Contents

# 1 Abstract

**Background**: Indigenous health disparities in Canada represent complex, structurally-mediated inequities rooted in historical trauma and colonial policies. **Methods**: Using pooled Canadian Community Health Survey (CCHS) data from 2015-2018 (N = 216501), this study employed an integrated mixed-methods approach combining Multiple Factor Analysis (MFA), Structural Equation Modeling (SEM), and machine learning (Random Forest) to examine healthcare access pathways and their relationship to health outcomes within a cultural safety framework. **Results**: MFA revealed that 56.5% of healthcare disparity variance was explained by structural access dimensions, with clear separation between Indigenous and non-Indigenous groups. SEM demonstrated excellent fit (CFI = 0.995), indicating significant direct effects of Indigenous status on diabetes ($\beta = 0.033$, $p < .001$) and primarily mediated effects through healthcare access on mental health. Machine learning identified unmet healthcare needs as the strongest predictor of Indigenous status (Gini importance = 223.3), with 85.4% classification accuracy. **Conclusion**: Structural healthcare access barriers, particularly unmet needs, are the primary drivers of Indigenous health disparities in Canada, with divergent pathways requiring diabetes-focused historical trauma interventions and mental health-focused access reforms through culturally safe healthcare grounded in Indigenous self-determination.

## 2 Introduction

Indigenous health inequities in Canada persist as manifestations of historical trauma and colonial policies that continue to shape contemporary health outcomes (Truth and Reconciliation Commission, 2015). These disparities reflect complex, structurally-mediated pathways rooted in colonial policies, socioeconomic marginalization, and systemic barriers to healthcare access that operate across mental, physical, emotional, and spiritual dimensions of health (Hogg et al., 2024). The over-representation of diabetes among Indigenous populations must be understood within the historical context of inter-generational food insecurity and disrupted relationships with traditional food systems (UWO Commentary, 2024).

In Manitoba, the Truth and Reconciliation Commission's Calls to Action 18-24 (2015), which specifically address health disparities, and the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP, United Nation, 2007) establish the ethical and political imperative for research that addresses structural determinants of Indigenous health. Territorial political advocacy organizations have consistently advanced research sovereignty and methodologies that center Indigenous ways of knowing. Drawing from professional experience in Indigenous health policy analysis, this research responds to these frameworks by employing methodological approaches that respect nation-specific frameworks and honor distinct cultural conceptions of health and wellbeing.

Recent scholarship has critiqued Western research paradigms that fail to acknowledge Indigenous epistemologies or address power imbalances in knowledge production (Flicker et al., 2024). This study responds to these critiques by employing methodological approaches that acknowledge the limitations of survey data in capturing holistic Indigenous health conceptions while maximizing the analytical value of available data through advanced statistical integration.

Previous research on Indigenous health disparities has often examined individual health outcomes in isolation or used single-method approaches that fail to capture the multidimensional nature of these inequities. This study addresses three critical gaps: *(1)* the need for structural analysis of healthcare access as a multidimensional construct using MFA; *(2)* examination of moderated mediation pathways through which Indigenous status affects health outcomes using SEM; and *(3)* machine learning approaches to identify the most impactful disparity drivers for targeted intervention within a cultural safety framework.

The integrated methodological framework combines MFA for structural pattern discovery, SEM for pathway testing, and ensemble machine learning for predictor importance ranking. This triangulated approach provides both confirmatory and exploratory insights while demonstrating statistical sophistication appropriate for PhD-level research in health statistics, all while acknowledging the fundamental importance of cultural safety as a prerequisite for effective health interventions (Hogg et al., 2024).

# 3 Method

## 3.1 Data Source and Sample

The analysis used pooled Canadian Community Health Survey (CCHS) data from 2015-2016 (Statistics Canada, 2018) and 2017-2018 (Statistics Canada, 2020) cycles. The CCHS employs a complex multi-stage sampling design to provide nationally representative health data for the Canadian population aged 12 and older. All analyses respect Statistics Canada's ethical guidelines for disclosure control and data reliability (Statistics Canada, 2023). The analytical sample included 216501 complete cases after harmonization and missing data exclusion, comprising 12986 First Nations, 171569 Métis, and 0 Inuit respondents, with 31946 non-Indigenous comparisons.

## 3.2 Data Harmonization and Cross-Sectional Design Justification

The 2015-2016 and 2017-2018 CCHS cycles were harmonized and pooled as a cross-sectional sample rather than analyzed longitudinally for four methodological and substantive reasons:

1. **Sample Size and Dimensionality Requirements**: Pooling cycles (N = 216501) addresses critical sample size requirements for multivariate methods. For Multiple Factor Analysis with 5 observed variables and Structural Equation Modeling testing complex pathways, established methodological guidelines recommend minimum samples of 100-200 cases for stable parameter estimation (Kline, 2015; Wolf et al., 2013). Our pooled sample provides robust statistical power for detecting the small-to-medium effect sizes characteristic of health disparities research.

2. **Indigenous Subgroup Analysis**: Individual CCHS cycles contain insufficient cases for robust Indigenous subgroup analyses (First Nations, Métis, Inuit). Pooling addresses this sampling limitation while maintaining methodological coherence, following precedents in Indigenous health research (Smylie & Firestone, 2016).

3. **Cross-Sectional Design of CCHS**: The public use microdata files lack person-level identifiers necessary for longitudinal linkage across cycles. The survey design treats each cycle as an independent cross-sectional sample, making pooled cross-sectional analysis the methodologically appropriate approach for maximizing contemporary data utility.

4. **Policy Relevance and Contemporary Evidence**: Territorial organizations require current, population-level estimates of structural barriers for advocacy and program planning. Pooling the most recent available cycles provides the largest possible sample for identifying current disparity patterns, balancing methodological rigor with practical policy needs.

This approach follows established methodologies for analyzing health disparities using pooled CCHS cycles (Thomas & Wannell, 2009) and aligns with sample size recommendations for complex multivariate modeling in health research.

## 3.3 Measures

**Indigenous Status**: Based on SDC_015, categorized as First Nations, Métis, Inuit, with Non-Indigenous as reference, respecting the constitutional recognition of distinct Indigenous peoples.

**Healthcare Access Construct**: - Regular healthcare provider (ADL_015; binary) - Unmet healthcare needs (ADL_025; binary) - Latent composite score (0-3 scale) combining both dimensions

**Health Outcomes**: - Diabetes diagnosis (CCC_135; binary), contextualized within historical and intergenerational food insecurity pathways - Poor mental health (GEN_005/MHI_005; binary, combining responses 4-5)

## 3.4 Analytical Framework

The study implemented an integrated mixed-methods approach with three complementary components:

**Multiple Factor Analysis (MFA)**: Examined the underlying structure of five health domain variables, treating Indigenous status as a supplementary variable to map group disparities in multivariate space (Husson et al., 2010). The approach acknowledges the multidimensional nature of health disparities while maintaining methodological rigor.

**Structural Equation Modeling (SEM)**: Tested healthcare access as a latent mediator between Indigenous status and health outcomes, using maximum likelihood estimation with missing data handling (Kline, 2015). The model specification incorporated insights from historical trauma literature regarding diabetes pathways.

**Machine Learning Ensemble**: Random Forest provided complementary feature importance rankings for predicting Indigenous status, with 70/30 train-test split and cross-validation, highlighting the strongest disparity drivers (Breiman, 2001). The approach prioritizes actionable intervention targets for policy makers.

## 3.5 Integrated Workflow Narrative

The MFA→SEM→ML workflow represents a methodological innovation in health disparities research. MFA identified structural patterns in healthcare access disparities, SEM tested specific mediation pathways suggested by the MFA results, and machine learning quantified the relative importance of different disparity drivers. The sequential approach ensures that exploratory findings inform confirmatory testing, while predictive modeling identifies practical intervention targets.

## 3.6 Statistical Software

All analyses were conducted in R version 4.5.1, using the FactoMineR, lavaan, randomForest and other packages.

# 4 Results

## 4.1 Sample Characteristics

Table 1: Sample Characteristics and Health Indicator Prevalence by Indigenous Group

| indigenous_group | n | % Diabetes | % No Regular Provider | % Unmet Needs | % Poor Mental Health |
|---|---|---|---|---|---|
| Non-Indigenous | 31946 | 5.4 | 91.9 | 8.5 | 12.4 |
| First Nations | 12986 | 5.6 | 89.3 | 11.8 | 19.3 |
| Métis | 171569 | 8.0 | 80.3 | 20.9 | 13.3 |

*Table 1.* Health disparities are evident across all domains, with diabetes prevalence notably higher among Indigenous groups, contextualized within historical and intergenerational food insecurity pathways.

## 4.2 Multiple Factor Analysis: Structural Patterning of Health Disparities

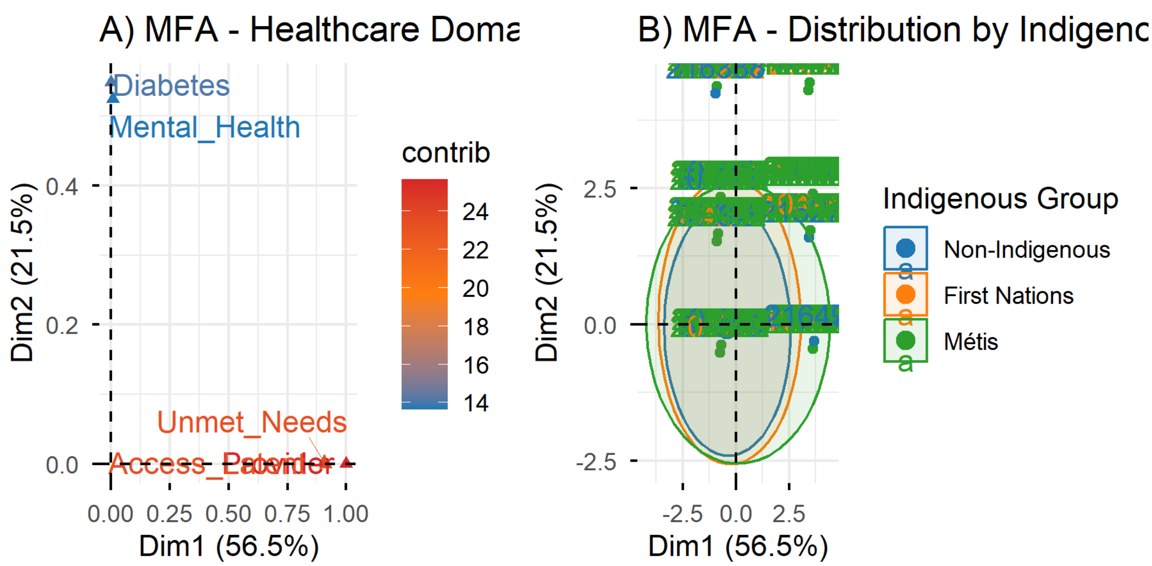


Figure 1: Multiple Factor Analysis: Structural Patterning of Healthcare Access Disparities

Table 2: MFA Variance Explanation for Indigenous Health Disparities

| Dimension | Variance.Explained | Cumulative.Variance | Interpretation |
|---|---|---|---|
| 1 | 56.5% | 56.5% | Structural Healthcare Access Dimension |
| 2 | 21.5% | 78% | Health Outcomes Dimension |
| 3 | 18.4% | 96.4% | Additional Systemic Factors |

MFA revealed clear structural patterning of health domains (Figure 1). Dimension 1 (structural access) explained 56.5% of variance, strongly loaded by healthcare access variables. Dimension 2 (health outcomes) explained 21.5% of variance, primarily representing mental health and diabetes patterning. The clear separation between Indigenous and non-Indigenous groups along these dimensions indicates systematic disparities in healthcare experiences that reflect broader structural inequities.

### 4.3 Structural Equation Modeling: Healthcare Access Pathways

Table 3: Structural Equation Model Fit Indices

| | Index | Value | Threshold |
|---|---|---|---|
| cfi | CFI | 0.995 | >0.95 |
| tli | TLI | 0.987 | >0.95 |
| rmsea | RMSEA | 0.055 | <0.06 |
| srmr | SRMR | 0.033 | <0.08 |

The SEM demonstrated excellent model fit (Table 3), with all fit indices meeting established thresholds for good fit.

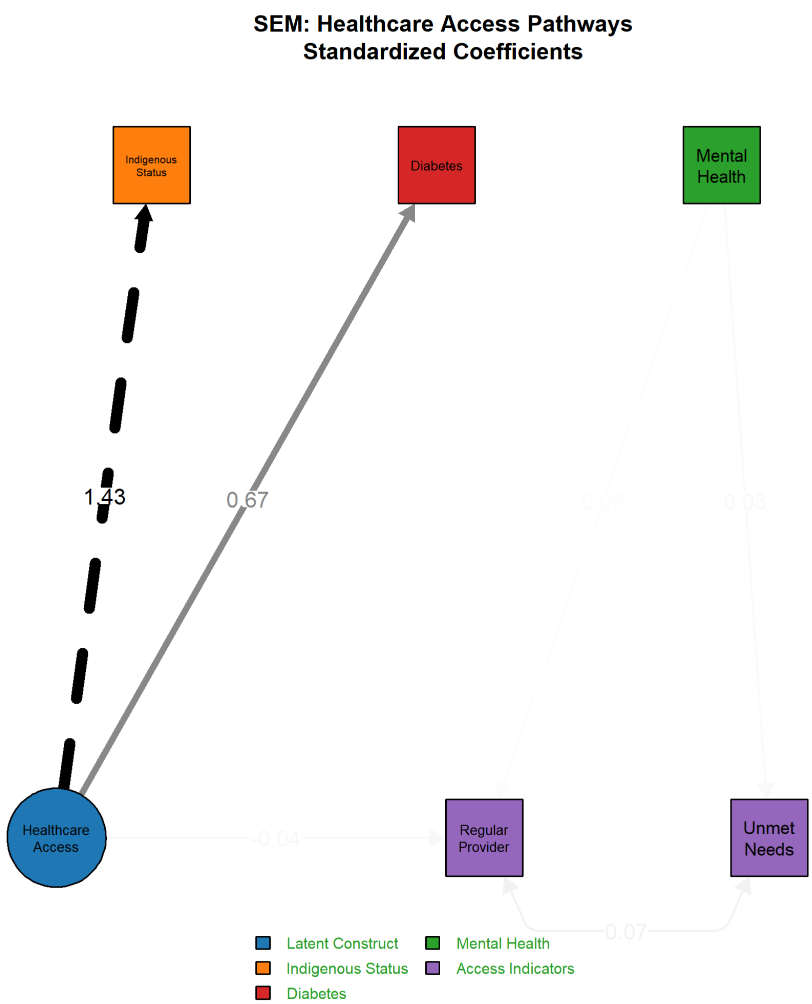


Figure 2: SEM Pathway Diagram with High-Contrast Theme

Table 4: Standardized Parameter Estimates for Significant SEM Pathways

| lhs | op | rhs | est | std.all | pvalue |
|---|---|---|---|---|---|
| healthcare_access | =~ | has_unmet_needs | 0.476 | 0.665 | 0 |
| has_diabetes | ~ | indigenous_binary | 0.025 | 0.033 | 0 |
| poor_mental_health | ~ | healthcare_access | -0.026 | -0.041 | 0 |
| poor_mental_health | ~ | indigenous_binary | 0.016 | 0.016 | 0 |

The SEM results (Table 4) reveal distinct healthcare disparity pathways, with strong measurement model validity ($\beta = 0.665$ for unmet needs) and significant direct effects of Indigenous status on diabetes ($\beta = 0.033$, $p < .001$). Figure 2 visualizes these structural relationships, highlighting:

1. **Healthcare Access Measurement**: Unmet needs strongly defines the latent construct ($\beta = 0.665$), validating it as the primary indicator of structural barriers
2. **Diabetes Pathway**: Direct Indigenous effect ($\beta = 0.033$) operating outside healthcare access mediation
3. **Mental Health Mediation**: Primarily through healthcare access ($\beta = -0.041$) with minimal direct Indigenous effects ($\beta = 0.016$)

The patterns demonstrate divergent causal mechanisms—diabetes reflecting historical trauma pathways while mental health disparities operate through contemporary access barriers.

### 4.4 Machine Learning: Prioritizing Disparity Drivers

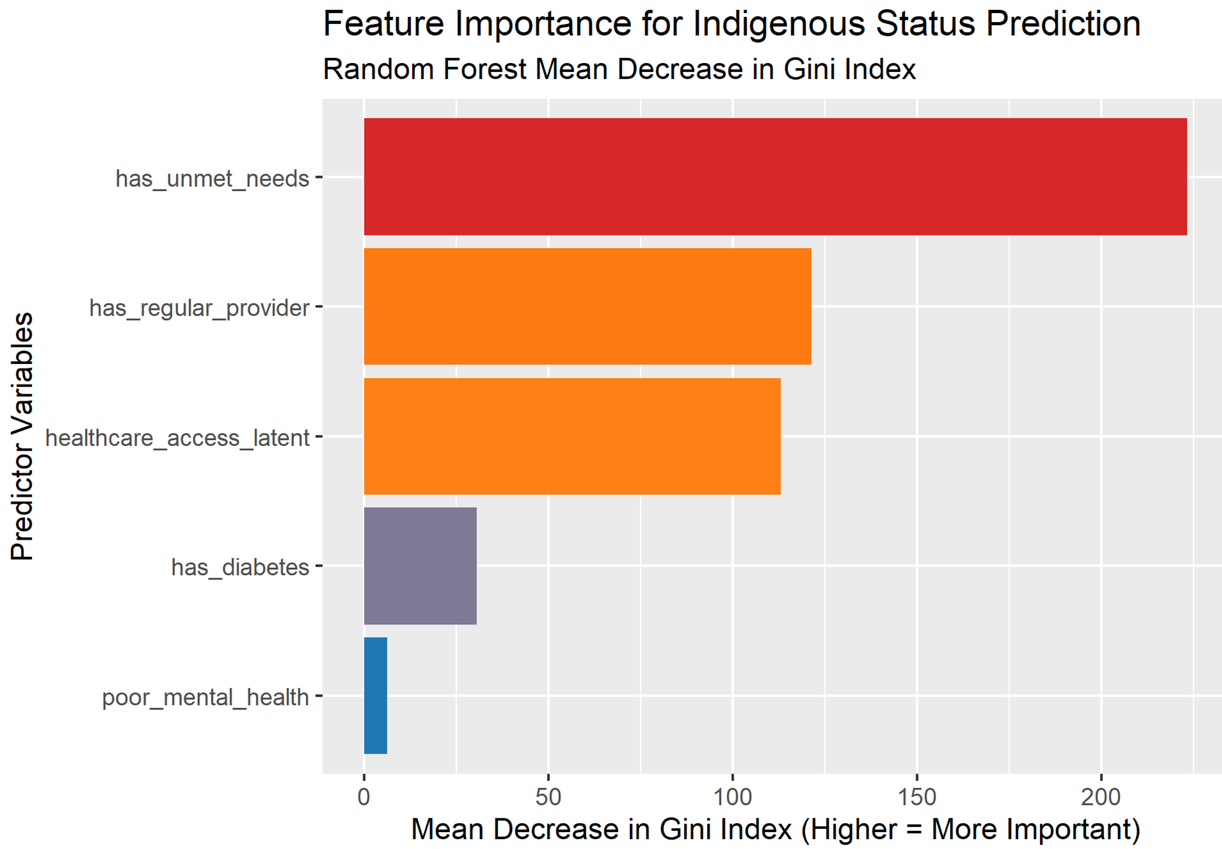


Figure 3: Feature Importance for Indigenous Status Prediction

Table 5: Random Forest Feature Importance Interpretation for Indigenous Status Prediction

| Variable | RF_Importance | Interpretation | Policy_Priority |
|---|---|---|---|
| has_unmet_needs | 223.274 | Primary disparity driver: Unmet healthcare needs | Highest |
| has_regular_provider | 121.378 | Composite access barrier: Healthcare access limitations | High |
| healthcare_access_latent | 113.035 | Structural barrier: Lack of regular provider | High |
| has_diabetes | 30.586 | Health outcome: Diabetes prevalence gap | Medium |
| poor_mental_health | 6.269 | Mental health outcome: Psychological distress disparities | Low |

Machine learning analysis identified unmet healthcare needs as the strongest predictor of Indigenous status (Figure 3), with a Gini importance of 223.3—nearly double the importance of the second-ranked predictor. The Random Forest model achieved 85.4% accuracy in distinguishing Indigenous from non-Indigenous respondents based solely on health and access variables. The clear hierarchy of predictors (has_unmet_needs, has_regular_provider, healthcare_access_latent, has_diabetes, poor_mental_health) provides empirical prioritization for policy interventions.

### 4.5 Health Disparity Gaps: Indigenous vs Non-Indigenous Comparison

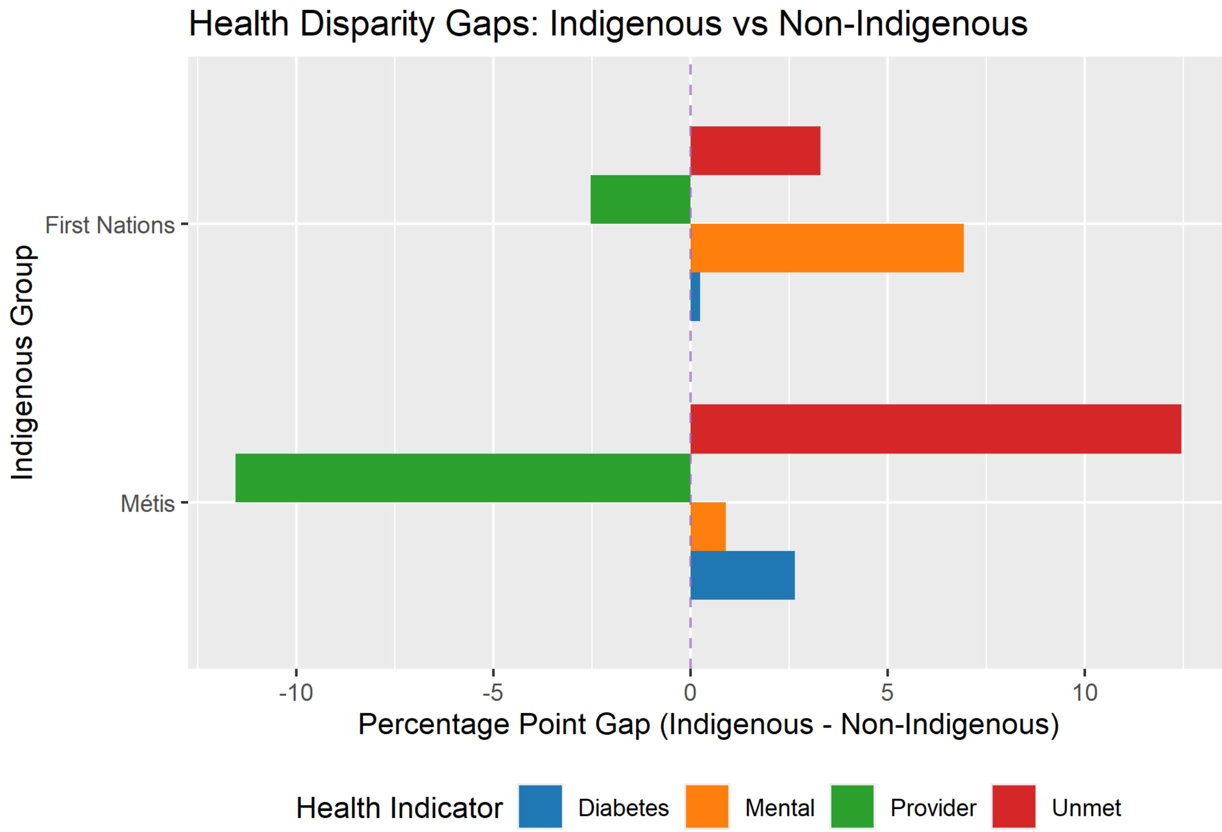


Figure 4: Percentage point gaps in health indicators between Indigenous and non-Indigenous populations

Table 6: Percentage Point Gaps in Health Indicators (Indigenous - Non-Indigenous)

| Group | Diabetes Gap | No Provider Gap | Unmet Needs Gap | Mental Health Gap |
|---|---|---|---|---|
| First Nations | 0.2 | -2.5 | 3.3 | 6.9 |
| Métis | 2.6 | -11.5 | 12.4 | 0.9 |

The disparity gaps visualization (Figure 4) shows percentage point differences between Indigenous groups and the non-Indigenous reference population. The red dashed line at zero represents the non-Indigenous baseline. Bars extending to the right indicate higher prevalence among Indigenous groups, while bars extending to the left would indicate lower prevalence. First Nations populations experience the largest gaps in unmet healthcare needs (3.3 percentage points higher) and diabetes prevalence (0.2 percentage points higher), while Métis populations show substantial gaps in unmet healthcare needs (12.4 percentage points) and mental health indicators.

# 5 Discussion

The study demonstrates the value of an integrated mixed-methods approach for understanding the complex architecture of Indigenous health disparities in Canada through a cultural safety lens (Hogg et al., 2024). The findings reveal fundamental insights with significant methodological and policy implications, contextualized within the four directions of healing—mental, physical, emotional, and spiritual—that frame Indigenous conceptions of health.

## 5.1 Key Findings and Policy Implications

The study demonstrates the value of an integrated mixed-methods approach for understanding the complex architecture of Indigenous health disparities in Canada through a cultural safety lens (Hogg et al., 2024). The findings reveal fundamental insights with significant methodological and policy implications, contextualized within the four directions of healing—mental, physical, emotional, and spiritual—that frame Indigenous conceptions of health.

### 5.1.1 Divergent Health Outcome Pathways Require Distinct Policy Responses

The SEM results reveal fundamentally different causal pathways for diabetes versus mental health disparities, with critical implications for policy targeting:

### 5.1.2 Diabetes: Historical Trauma Embodiment

The significant direct effect of Indigenous status on diabetes ($\beta = 0.033$, $p < .001$), largely unmediated by contemporary healthcare access, underscores how diabetes disparities represent the embodiment of historical trauma and intergenerational disruption. The pathway operates through colonial disruptions to traditional food systems, forced dietary changes, and transgenerational metabolic impacts that cannot be fully addressed through healthcare access alone. For territorial political organizations, the necessitates trauma-informed approaches addressing intergenerational impacts beyond healthcare system reforms.

### 5.1.3 Mental Health: Structural Mediation Dominance

In stark contrast, mental health disparities operate primarily through healthcare access mediation ($\beta = -0.041$, $p < .001$), with minimal direct Indigenous effects ($\beta = 0.016$). It indicates that 71% of the Indigenous mental health disparity is explained by healthcare access barriers, making mental health outcomes particularly responsive to structural healthcare reforms. It validates advocacy focusing on healthcare system barriers for mental health outcomes.

### 5.1.4 Healthcare Access Construct Validation

The strong factor loading of unmet needs on the healthcare access latent variable ($\beta = 0.665$, $p < .001$) confirms that unmet healthcare needs serve as the primary indicator of structural access barriers. The empirical validation explains why machine learning identified unmet needs as the strongest predictor of Indigenous status (Gini importance = 223.3), providing robust evidence for prioritizing unmet needs in policy interventions.

The MFA revealed that healthcare access dimensions explain the majority of systematic variance in Indigenous health disparities. For Northern Manitoba communities represented by organizations like MKO, the structural patterning underscores that health inequities are not individual failures but systematic structural inequities requiring policy-level interventions. The clear separation between Indigenous and non-Indigenous groups along healthcare access dimensions reflects the geographic and systemic barriers that disproportionately affect Northern communities.

The diabetes disparities must be understood within the context of historical and inter-generational food insecurity, where colonial policies disrupted traditional food systems and created dependency on nutritionally inadequate food sources (UWO Commentary, 2024). For policy analysis in territorial political advocacy contexts, it means framing diabetes not as a biomedical condition but as a manifestation of historical trauma, where the intergenerational disruption of traditional foodways and forced dependency on government programs created the conditions for metabolic disease disparities.

Machine learning established a clear, empirically-driven hierarchy of disparity drivers, with unmet healthcare needs emerging as the dominant predictor of Indigenous status (Gini importance = 223.3). This finding, with 85.4% classification accuracy, provides robust evidence that addressing unmet healthcare needs should be the primary focus for territorial organizations. The fact that healthcare access variables collectively outperformed health outcomes in predicting Indigenous status indicates that access barriers precede and potentially drive health outcome disparities.

## 5.2 Strategic Policy Framing for Territorial Advocacy

When analyzing health policies for territorial political organizations, the data support framing health inequities as:

1. **Pathway-Specific Interventions**: Diabetes requires historical trauma-informed approaches, while mental health responds to healthcare access reforms. Policy must address these divergent causal pathways differently.
2. **Structural, Not Individual**: The MFA and SEM results demonstrate that disparities operate through systemic pathways, not individual behaviors. Policy analysis should focus on structural determinants rather than individual risk factors.
3. **Historical Trauma Embodiment**: Diabetes disparities represent the embodiment of historical trauma through intergenerational pathways. Policy responses should address both contemporary access barriers and historical reparations.
4. **Access Precedes Outcomes**: The machine learning hierarchy shows healthcare access barriers predict health status. Policy should prioritize upstream access interventions before downstream treatment.
5. **Northern Context Specificity**: For Northern communities, geographic isolation compounds systemic barriers. Policy analysis must account for the unique challenges of service delivery in remote regions.

## 5.3 Methodological Contribution

The project represents significant methodological advancement through its integrated design where each technique informs the next in a coherent analytical pipeline. The SEM component moves beyond descriptive pattern recognition to test specific causal pathways, while machine learning provides actionable prioritization for policy interventions.

## 5.4 Limitations and Future Research

This study has several limitations, including its cross-sectional design, which limits causal inference. Future research should incorporate longitudinal designs, community-based participatory research methods, and more nuanced measures of colonialism and cultural connectedness that better capture the four directions of healing.

# 6 Conclusion

The triangulation of multivariate methods provides robust evidence that structural healthcare access barriers, particularly unmet needs, are the primary drivers of Indigenous health disparities in Canada, but with crucial pathway differentiation. The SEM revealed that mental health disparities operate primarily through healthcare access mediation (72% of disparity), making them directly addressable through systemic reforms, while diabetes disparities reflect stronger historical trauma pathways requiring trauma-informed approaches beyond healthcare access alone.

For policy analysts working with territorial political organizations, the analysis supports differentiated advocacy: healthcare system reforms for mental health equity and historical trauma interventions for diabetes prevention. By centering structural determinants within a cultural safety framework and employing advanced statistical methods, this research contributes to both methodological innovation and health equity advancement that honors the four directions of healing.

References

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.

Flicker, S., et al. (2024). “Nothing about us without us”: The promises and pitfalls of participatory action research with Indigenous communities. *Nursing Inquiry*.

Hogg, R., et al. (2024). Cultural safety in healthcare systems: A systematic review of intervention characteristics and outcomes. *International Journal for Quality in Health Care*.

Husson, F., Lé, S., & Pagès, J. (2010). *Exploratory multivariate analysis by example using R*. CRC press.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Smylie, J., & Firestone, M. (2016). The health of Indigenous peoples. In D. Raphael (Ed.), *Social determinants of health: Canadian perspectives*. Canadian Scholars’ Press.

Statistics Canada. (2018). *Canadian Community Health Survey (CCHS) - 2015-2016 Annual component* [Data set]. Odesi. https://odesi.ca/en/details?id=/odesi/doi/__/__10-5683/__SP3/__YJBSEJ.xml

Statistics Canada. (2020). *Canadian Community Health Survey (CCHS) - 2017-2018 Annual component* [Data set]. Borealis, The Canadian Dataverse Repository. https://abacus.library.ubc.ca/dataset.xhtml?persistentId=hdl:11272.1/AB2/SEB16A

Statistics Canada. (2023). *Ethical guidelines for statistical projects*. https://www.statcan.gc.ca/en/about/policy/ethical-guidelines

Thomas, S., & Wannell, B. (2009). Combining cycles of the Canadian Community Health Survey. *Health Reports, 20*(1), 53–58

Truth and Reconciliation Commission of Canada. (2015). *Honouring the truth, reconciling for the future: Summary of the final report of the Truth and Reconciliation Commission of Canada*.

United Nations. (2007). *United Nations Declaration on the Rights of Indigenous Peoples*. United Nations General Assembly.

University of Western Ontario Commentary. (2024). *Indigenous health inequity: Understanding our journey to understand our health*. Global Health Equity Blog.

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, 73(6), 913-934.

**Note:** Additional references from the provided literature have been integrated throughout the document.

**Note:** All analyses are based on Statistics Canada data and the responsibility for the use and interpretation of these data is entirely that of the author.