

# Introduction to the Canadian Community Health Survey

With a focus of use at ICES to support studies examining equity issues

Doug Manuel

2025-06-02



The Ottawa  
Hospital  
Research Institute

L'Hôpital  
d'Ottawa  
Institut de recherche

**Inspired** by research.  
**Driven** by compassion.

**Inspiré** par la recherche.  
**Guidé** par la compassion.

Affiliated with Affilié à

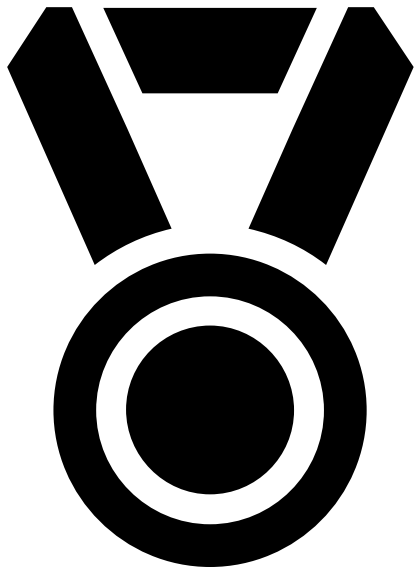


uOttawa

# Agenda

- Overview of the survey
- Getting started using the survey in your studies
- Examples of studies

# A notable population health survey



The largest ongoing population health survey worldwide.

Best-in-class design and implementation

National linkages to hospitalization and death, with improving linkages and analysis environment (RDC).

In ICES, more linkages to health data than any other survey worldwide

Public Use Microfile (PUMF) with Statscan Open License.

# But with challenges



Complex survey design – stratified multi-stage design



Several major redevelopments over the past 2+ decades (in 2015 and 2023).



1000 variables that regularly change over time.

# Successful approach



Follow established approaches (previous studies).



Use collaboration resource (and contribute to them). Yeah, Open Science!



Be diligent with data cleaning and preparation.

# **Part 1**

**Target population, survey  
design, selection, sample size,  
response rate,  
content**

# Target population

- Household-dwelling Canadians.
- Biannual from 2001 to 2007, then annual.
- Population aged 12 years and older.
  - Starting 2020, age increased 18 years and older in 2023.
- 10 provinces and three territories.
  - Territories stopped/reduced collection in 2020 and 2021.

# Exclusions - approximately 2–3% of the population

- People **living on reserves and other Aboriginal settlements** in the provinces.
- **Full-time members of the Canadian Forces.**
- **Institutionalized populations** (e.g., long-term care facilities, prisons).
- Some **remote areas** of the territories with very small populations or accessibility challenges.



# CCHS 2009–2010 Sampling Design

- **Three-step sample allocation:**
  - Minimum 500 respondents per Health Region (HR)
  - Remainder allocated by provincial population
  - Provincial sample split across HRs by  $\sqrt{(\text{HR population})}$
- **Three sampling frames:**
  - **Area frame (49.5%)** – LFS-based, PPS, adjusted for yield
  - **List frame (49.5%)** – Canada Phone Directory
  - **RDD (1%)** – Elimination of Non-Working Banks method
- **One person per household**, with oversampling of youth (12–19) and young adults (20–29)
- **Sample deployed evenly** over 2 years, in 2-month waves, controlling for seasonality

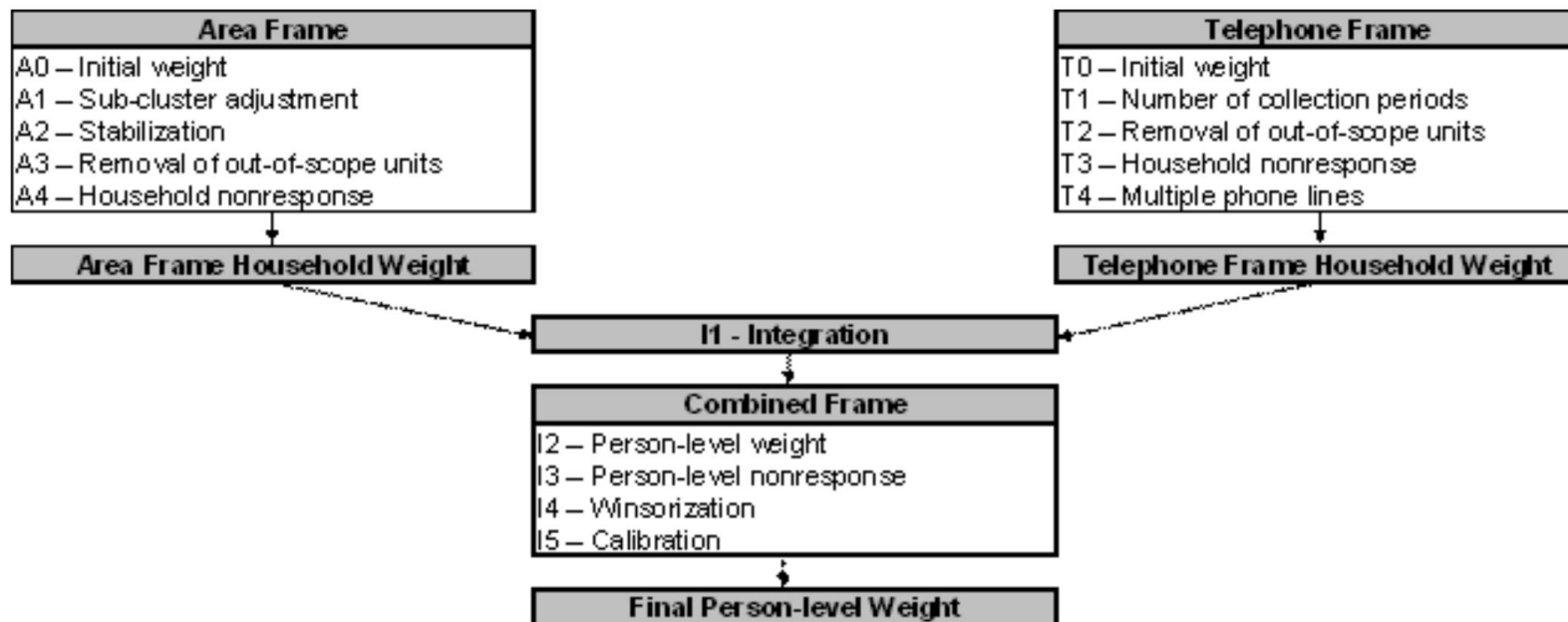
# Number of region and target sample size by province and territory, 2009-10

Province	Number of HRs	Targeted sample size 2009–2010
Newfoundland and Labrador	4	4,010
Prince Edward Island	3	2,002
Nova Scotia	6	5,041
New Brunswick	7	5,150
Quebec	16	24,289
Ontario <sup>1</sup>	36	44,379
Manitoba	10	7,500
Saskatchewan	11	7,720
Alberta	9	12,200
British Columbia	16	16,095
Yukon	1	1,200
Northwest Territories	1	1,200
Nunavut	1	700
Canada	121	131,486

# CCHS RESPONSE RATES

SURVEY	DATE RANGE	ONTARIO			SUBSET OF ONTARIO RESPONDENTS & AGREE TO SHARE/LINK	RESPONSE RATE OF LINKED ONTARIO
		TARGET	RESPONSE	RESPONSE RATE		
2000-2001	SEPT2000 TO NOV2001	47,900	39,278	82.0%	83.6%	68.6%
2003-2004	JAN2003 TO JAN2004	53,834	42,260	78.5%	79.7%	62.6%
2005-2006	JAN2005 TO JAN2006	54,883	42,260	77.0%	79.0%	60.9%
2007-2008	JAN2007 TO DEC2008	60,408	44,460	73.6%	77.9%	57.3%
2009-2010	JAN2009 TO DEC2010	64,598	44,379	68.7%	76.0%	52.2%
2011-2012	JAN2011 TO DEC2012	67,267	44,396	66.0%	73.5%	48.5%
2013-2014	JAN2013 TO DEC2014	70,756	44,576	63.0%	74.1%	46.7%

# Weighting strategy



# Pop quiz #1

- Which prevalence estimate is higher?
  - The weighted Canadian smoking prevalence.
  - The unweighted Canadian smoking prevalence.

# Resources

- A good public collection of CCHS documentation is here.

<https://osf.io/hkuy3/>

- Survey methods and analysis.

<https://statmodeling.stat.columbia.edu/2025/06/01/survey-statistics-it-is-the-people/>



# Pop quiz #1

- Was the CCHS designed for data linkage studies?

# Content (~ 1000 variables/cycle)

- Sociodemographic and economic
- Health status and chronic conditions
- Health behaviours
- Health care use
- Community & support
- Optional content (by region)



# Sociodemographic

## 1. Education

- Highest level of education attained

## 2. Income and material deprivation

- Total household income (grouped ranges)
- Main source of income
- **Food security** (Household Food Security Survey Module)

## 3. Employment

- Labour force status (employed/unemployed/out of labour force)
- Main job type (e.g., class of worker, industry, occupation)
- Job stress (in optional content for some provinces)

## 4. Housing and household composition

- Number of people in household
- Relationship to household members
- Tenure (own vs. rent)
- Type of dwelling (e.g., detached, apartment)

## 5. Demographic context

- Age, sex, marital status
- Immigration status (whether immigrant, and years since immigration)

# Community and support

## 1. Social support

- Perceived availability of support in times of need
- (e.g., having someone to confide in, count on for help)

## 2. Work-related stress

- Self-reported work stress (included in optional content for some regions)

## 3. Caregiving

- Provided unpaid help or personal care to family/friends due to long-term condition, disability, or aging

## 4. Life satisfaction and psychosocial well-being

- Overall life satisfaction
- Sense of control over decisions affecting life
- Personal impact of chronic illness

## 5. Community and neighbourhood

- Sense of belonging to local community
- Neighbourhood safety (perceived safety walking alone after dark)

## 6. Access to care

- Unmet health needs and their reasons

# Tools to help your analyses

## - part 2

# <https://613apps.ca/data-dictionary-builder/>

Choose survey

CCHS


Filter options

☒ Group variables

☐ French labels


Build dictionary

Data dictionary builder v1.3





Explore About

The **Data Dictionary Builder** app is developed and maintained by Dr. Michelle Guerrero and Joel Barnes of the Health Analysis Division, Statistics Canada.




Michelle Guerrero, Ph.D.  
Analyst





Joel Barnes, M.Sc.  
Analyst



# <https://big-life-lab.github.io/cchsfLOW/>

cchsfLOW 2.1.0

Get started

Articles ▾

Reference

Changelog

Search...

## cchsfLOW

*cchsfLOW* supports the use of the Canadian Community Health Survey (CCHS) by transforming variables from each cycle into harmonized, consistent versions that span survey cycles (currently, 2001 to 2018).


The CCHS is a population-based cross-sectional survey of Canadians that has been administered every two years since 2001. There are approximately 130,000 respondents per cycle. Studies use multiple CCHS cycles to examine trends over time and increase sample size to examine sub-groups that are too small to examine in a single cycle.

The CCHS is one of the largest and most robust ongoing population health surveys worldwide. The CCHS, administered by Statistics Canada, is Canada's main general population health survey. Information about the survey is found [here](#). The CCHS has a [Statistic Canada Open Licence](#).

### Concept

Each cycle of the CCHS contains over 1000 variables that cover the four main topics: sociodemographic measures, health behaviours, health status and health care use. The *seemingly* consistent questions across CCHS cycles entice you to combine them together to increase sample size; however, you soon realize a challenge...

Imagine you want to use BMI (body mass index) for a study that spans CCHS 2001 to 2018. BMI *seems* like a straightforward measure that is routinely-collected worldwide. Indeed, BMI is included in all CCHS cycles. You examine the documentation and find the variable `HWTAGBMI` in the CCHS 2001 corresponds to body mass index, but that in other cycles, the variable name changes to `HWTCGBMI`, `HWTDGBMI`, `HWTEGBMI`, etc. On reading the documentation, you notice that some cycles round the value to one decimal, whereas other cycles round to two digits. Furthermore, some cycles don't calculate BMI for respondents < age 20 or > 64 years. Also, some cycles



### Links

- [View on CRAN](#)
- [Browse source code](#)
- [Report a bug](#)
- [Calculators](#)

### License

- [MIT](#) + file [LICENSE](#)

### Community

- [Contributing guide](#)
- [Code of conduct](#)

### Citation

- [Citing cchsfLOW](#)

### Developers

- Doug Manuel  
Author, copyright holder
- Warsame Yusuf  
Author
- Rostyslav Vyuha

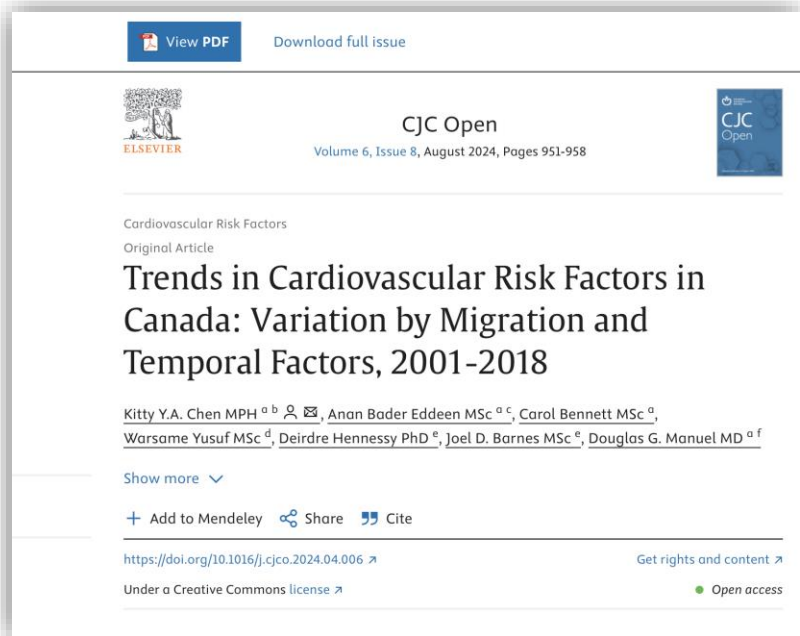


Discourse: <https://recode.discourse.group>

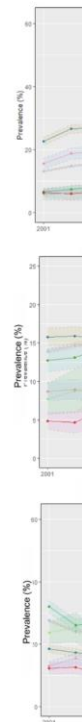
# Example

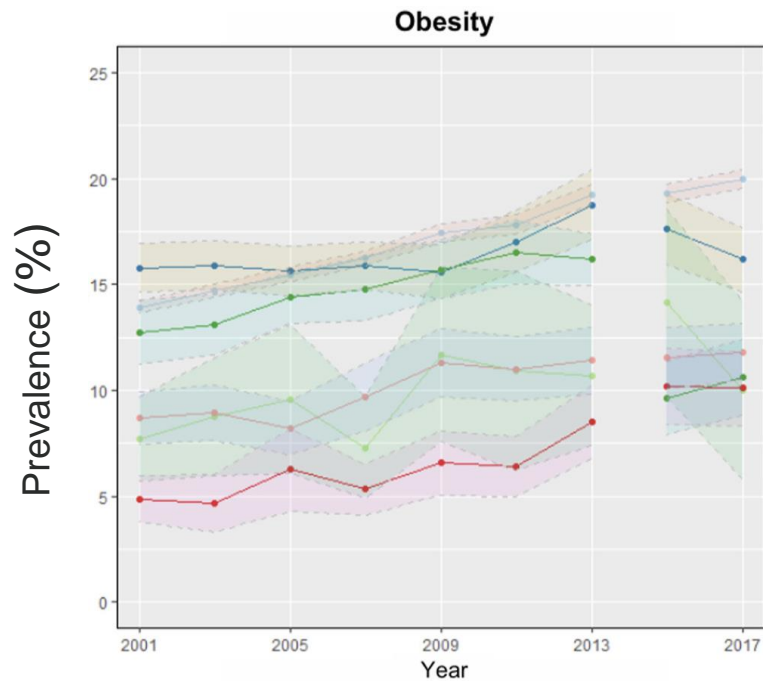
## - part 3



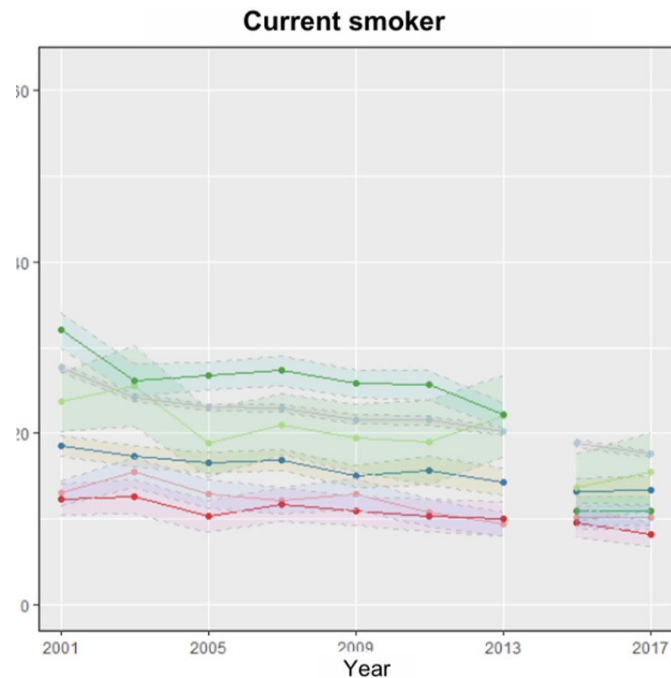


<https://github.com/Big-Life-Lab/cvd-trends-Canada>





- Non-racialized Canadian-born
- Racialized Canadian-born
- Non-racialized established immigrant



- Non-racialized recent immigrant
- Racialized recent immigrant



**Thank you**

A decorative graphic on the right side of the slide. It consists of several concentric circles of varying line weights. Superimposed on these circles are approximately 15 plus signs (+) of a light blue color. The plus signs are distributed across the circles, with some appearing on the outermost circle and others on the inner ones, creating a subtle geometric pattern.