**Preliminary Form to Add Social Determinants to CSDUL**

**Request date (2024-09-06):**

| **Researcher (name and affiliation):** | Anousheh Marouzi, Saskatchewan Health Authority |
| --- | --- |
| **Node Lead (name and affiliation):** | Charles Plante, Saskatchewan Health Authority |
| **Indicator or Model Name:** | Substance Use Harm Percentage |

**Purpose of the document**

This document includes several questions that must be answered by the researcher interested in adding indicators or models into CSDUL. These questions pretend to briefly explain the mathematical and theoretical framework of the indicator or model being incorporated. The researcher must be able to fill out every question clearly and concisely, supporting their explanation with respectable academic sources.

The document will be added to the model or indicator documentation in CSDUL-OUT and CSDUL-RDC. It must serve as a quick and straightforward introduction to the indicator or model for anyone interested and give relevant references to guide the learning process to other researchers.

**To be completed by the responsible analyst.**

**If there are questions that cannot be answered because of the nature of the indicator/model, write N/A.**

**You can support your completion using the example document located in this link:** [**Documents - Add inputs to CSDUL - 02 - Example.docx - Google Docs**](https://docs.google.com/document/d/1t4_Bh5pRtHzd8GQ3ifJWY2zjjjch8DFf/edit)

1. **Will you share the inputs through CSDUL-RDC, CSDUL-OUT, or both?** CSDUL-OUT.

1. **Explanation of the indicator/model.** 
   1. **In simple words, explain what the indicator/model to be added consists of.**

The indicator is the percentage of people with different sociodemographic characteristics. This percentage is calculated for the whole sample population, and two subpopulations, including people that have experienced substance use harm (PESUH) between 2006 and 2016, and people who have not experienced substance use harm (NPESUH) during this time.

* 1. **Are there assumptions associated with the indicator/model? If there are, please briefly describe them.** N/A.
  2. **How is the indicator/model derived? Support your explanation with formulas when possible.** The indicators are percentages among all cohort members, PESUH, and NPESUH. That is, the denominator is the whole population in the cohort, PESUH population, and NPESUH population, respectively. These percentages are then calculated for each sociodemographic variable. For example for male PESUH:
  3. **What is the unit of analysis of the indicator/model? (e.g. households, persons, cities)**

Saskatchewan province by different sociodemographic characteristics, including sex, after-tax household income quintile, rurality, ethnicity, occupation, employment, and education.

* 1. **How can the indicator be integrated with other datasets?**

The indicator can be merged with other datasets based on Saskatchewan and different values of the abovementioned sociodemographic variables.

* 1. **What are the boundaries of the indicator/model?**

0 to 100.

* 1. **If you want to add a model to CSDUL, is this associated with a hypothesis? If yes, please describe their:**
     1. **Null hypothesis**
     2. **Alternative hypothesis**
     3. **The implications of rejecting the null hypothesis**

**N/A**

* 1. **What is the interpretation of the values of the indicator/model?**

The percentage shows how the PESUH and NPESUH are distributed across different population groups in terms of different sociodemographic characteristics.

* 1. **Based on the literature and your experience working with this indicator/model, is it possible to identify weaknesses in its calculations or assumptions? To facilitate your answer, you can focus on:**

1. **Potential biases**
2. **Overestimation**
3. **Underestimation**
4. **Omitted variables**
5. **Endogeneity**
6. **Datasets’ problems**

The data source used to calculate the percentages is CanCHEC. Estimates produced using this database are not generalizable to the whole Canadian population. Since CanCHEC excludes the institutional population[[1]](#footnote-0) (e.g., those living in nursing homes, penitentiaries, group homes) at baseline,[[15]](https://paperpile.com/c/rqu83V/hHYhj) the created cohort is representative of the non-institutional population living in Canada at the time of the census. This makes the cohort population younger and healthier than the Canadian population.

Our cohort of Saskatchewanians might include individuals with hospitalization or mortality records over those without. This is due to our approach regarding identifying Saskatchewanians, where in addition to the census, we used DAD (2006-2016) and CVSD (2006-2016) data to find people who had an indication of residency in Saskatchewan.

1. **Does the indicator/model have other mathematical or computational versions (not syntax) to build it? (provide references)** While the percentage itself is calculated in only one way, the identification of PESUH can be done in various ways. For example, the Canadian Institure for Health Information (CIHI)[(CIHI 2022)](https://paperpile.com/c/FPUDjD/K0bV) and Canadian Substance Use Costs and Harms Scientific Working Group have developed their own substance use case-finding algorithm[(Canadian Substance Use Costs and Harm...)](https://paperpile.com/c/FPUDjD/kXnX). Adapting different case-finding algorithms would potentially alter the calculated percentages.
   1. **Why are you building the indicator/model as you propose? Are there advantages compared to other versions?** We used the 10th revision of the International Statistical Classification of Diseases and Related Health program (ICD-10)[[24]](https://paperpile.com/c/rqu83V/PgXq7) codes to identify hospitalizations or deaths that happened due to substance use. The case-finding algorithm used to identify SUH events is provided in Appendix 3 of the article titled “Creating an 11-year longitudinal substance use harm cohort from linked health and census data to analyze social drivers of health” submitted as a supporting file. This algorithm was developed based on methodologies applied by other researchers.[[12,25–28]](https://paperpile.com/c/rqu83V/fTAZs+NZ6dd+oD5Cq+PEpA3+IWpzE) We did not include opioid-related adverse drug reactions (Y45.0) as we were not interested in the harms resulting from the adverse effects of prescribed medications. It should also be noted that this list only contains harms that are 100% attributable to substance use, and therefore harms partially attributable to substance use (e.g., cancer, stroke) are excluded from this analysis.
2. **Do you see potential improvements for the indicator/model? This could involve using other datasets, refining calculations, or modifying assumptions, among others.**

**This indicator could be improved by including other datasets, such as emergency data, to identify more substance use cases. The indicator can also be extended to cover the whole Canadian population and can be broken down into smaller geographies like CMAs. This can be done by modifying the Stata do-file submitted as a supporting file.**

**What inputs are to be added to CSDUL? Write “X”**

|  | Raw or intermediate datasets required to create the indicator/model. |
| --- | --- |
| **X** | Codes that create the indicator/model (be sure that your code is clear enough to be replicated in the future for yourself or any other researcher). |
| **X** | Documentation that explains step by step the entire process that builds the indicator or model. |
| **X** | Results, which consist of the list of variables, indicators, or model results. |
|  | Support files. They can be papers, chapter books, codes, etc. |

**References**

1. CancCHEC considers individuals institutionalized if they do not have any other residency address in Canada and they have been in an institution for not less than six months. [↑](#footnote-ref-0)