

The impact of the POP-Grouper reference period on the results of the segmentation of clinical profiles from clinical-administrative data from Quebec.

**Authors:** Elhadji Malick Ndao (a), Mamadou Diop (a), Mike Benigeri (a), Jean-Luc Kaboré (a), Isabelle Samson (a)

## Introduction

The effective organization of health services is based on a precise and up-to-date segmentation of the clinical profiles of the population. For example, the Canadian Institute for Health Information's (CIHI) POP-Grouper was adapted to Quebec data as a tool for model adjustments and service utilization projections. However, the results of segmentation are highly dependent on the reference period for the identification of health conditions. A longer period of time allows for more conditions to be captured, although some, such as acute conditions, may no longer reflect the patient's clinical profile for a given year. This study aims to evaluate the impact of the reference period on segmentation in order to determine the best approach to obtain an accurate representation of clinical profiles.

## Methods

The POP-Grouper tool classifies individuals according to their diagnoses recorded in the clinical-administrative databases, into 226 conditions and 16 clinical categories.

The study compares segmentations over several periods, for example one year versus three years. The analysis is based on measures of concordance and variation based on performance models including new learning techniques, in order to assess compliance in terms of sensitivity and specificity. The results will be used to determine the reference period that maximizes the robustness and accuracy of the segmentation.

In addition, expert opinions will be collected to validate the relevance of the selected periods and to explore the combined use of two distinct periods. This process aims to ensure that segmentation is clinically relevant and suitable for different uses.

## Results

The POP-Grouper application on the Quebec population estimated at 8.8 million in 2023-2024 reveals that the use of a reference period of one year, rather than three years, significantly increases the number of people classified as nonusers of care or users of care without a condition, from 2.2 to 4.5 million people. This change from a one-year period to three years mainly impacts individuals with chronic conditions, indicating that the use of a longer period would allow for better identification of chronic conditions.

However, by moving from a three-year period to one year, a significant proportion of people in certain categories, such as newborns and obstetrics, are reclassified as non-users or users without affection (29%) or acute (26%), which better reflects the annual number of births. Similarly, 42% of people with acute conditions become non-users or users without affection, suggesting that a three-year period would overestimate these conditions.

## Discussion/Conclusions

These results show that the choice of the reference period has a major impact on the results of the segmentation based on clinical-administrative data. A long period facilitates the identification of chronic conditions, while a short period limits the overestimation of acute conditions. Finding the optimal method is essential to ensure relevant segmentation adapted to a good assessment of the health needs of populations. In addition, the reliability and completeness of the diagnoses in the databases are other issues to be considered.

a: Institut national d'excellence en santé et en services sociaux (INESSS), Canada