

Ascertaining the Francophone population in Ontario, validating the language variable in health data

OBJECTIVE and BACKGROUND:

Having good linguistic data is important for studying health inequalities within minority language groups. However, the validity and reliability of linguistic data from administrative health databases used in Canada have never been studied. Therefore, this study aims to identify which databases—those for home care, long-term care, and mental health admissions—best match linguistic data from the Canadian Community Health Survey (CCHS), a national survey that includes linguistic groups in Ontario. Since several existing linguistic variables exist to define and describe linguistic groups, the study chose the variables “home language” (HLS) and “first official language spoken” (FOLS) for comparison, as they are among the most commonly used in Canada.

KEY RESULTS & INTERPRETATIONS :

- The administrative databases for home care and long-term care are most consistent with the CCHS results for the LPM and FOLS, respectively.
 - Healthcare professionals providing these services are expected to observe the patient to determine if they need an interpreter, which may explain why these data correspond best.
 - The specificity of the results was high for both language variables, but the sensitivity for FOLS was lower. This could be because this variable includes mother tongue in its definition, and patients who no longer use French at home would not have been identified as Francophone in these databases.
- The database that least closely matches the CCHS results is the mental health admissions database.
 - This could be due to the fact that the care provided in these facilities is emergency care; therefore, they would have little time to properly assess language. Additionally, patients visiting these facilities require mental health care, which could also affect data collection if they are unable to provide valid information.

POLICY CONSIDERATIONS:

The results of this study do not imply a single approach to identifying language groups. Selecting the most appropriate language variable for a study should be guided by the study design and research question. However, these results may indicate the need to:

Improve language data collection

- Implement standardized procedures across all health databases to ensure accurate language data collection.
- Train healthcare professionals on the importance of this collection, including in urgent care settings such as mental health admissions.
- Implement regular data quality monitoring to identify areas for improvement.
- Integrate performance indicators related to language data quality into contracts or agreements with healthcare providers.

Strengthen access to services for linguistic minorities

- Ensure that Francophone patients are correctly identified so they can receive services in their language, particularly in areas where services are limited.
- Use reliable linguistic data to plan the provision of French-language health services and reduce inequalities in access.

For more information, see: Batista, R., Hsu, A., Bouchard, L. et al. Ascertaining the Francophone population in Ontario: validating the language variable in health data. *BMC Med Res Methodol* 24, 98 (2024). <https://doi.org/10.1186/s12874-024-02220-7>

For questions or comments, please contact Ricardo Batista (rbatista@ohri.ca).

This brief was developed in partnership with the Ontario SPOR Support Unit Francophone Initiative. Visit us at <https://savoirmontfort.ca/initiative-francophone/>

This text was written by the author, with support from OpenAI's ChatGPT tool for style improvement and suggestions for policy considerations based on research findings.