

Linking Hospital and Tax data to support research on the economic impacts of hospitalization

Sanmartin, Claudia¹, Reicker, Alexander¹, Garland, Allan², Iwashyna, Theodore³, Fransoo, Randy², Scales, Damon⁴, Wunsch, Hannah⁴, Forget, Evelyn², and Qiu, Hanqing¹

¹Statistics Canada

²University of Manitoba

³University of Michigan

⁴Sunnybrook Hospital

Objectives

This project links data on acute inpatient hospitalizations from the Canadian Discharge Abstract Database (DAD) with data on income and employment from various taxation- and employment-based administrative files. The goal was to create a linked database that will support research on the labour market and financial outcomes experienced by individuals and families following acute illness requiring hospitalization.

Approach

Data from the 1999/00 to 2014/15 Discharge Abstract Database (DAD) were linked to the 1981-2013/14 T1 Tax filer data and the Canadian Child Tax Benefit data. We sought to create a unique association between Health Insurance Numbers (HIN) available in the DAD and Social Insurance Numbers (SIN) available in the tax data by using variables common to both data sets - date of birth, postal code and sex. Both transactional data sets were "individualized" such that unique combinations of the linkage variables were identified and eligible for linkage. The linkage was conducted using deterministic methods.

Results

Approximately 97% of combinations involving date of birth, postal code and sex in the hospitalization data were uniquely related to a single valid HIN (n=18.8 million). Similarly, approximately 96% of the keys on the Tax data file were associated with a unique person. Approximately 86% of HINs were associated with a unique identifier in the tax file and these HINs

account for approximately 83% of the hospital records. The linkage was consistent over time, with linkage rates between 85% and 88% of HINs for all years. Some variation in linkage rates were observed by jurisdiction and by age. (Error estimates to be reported)

Conclusion

This project has created a unique linked database that will support research on the economic consequences of 'health shocks' for individuals and their families, and the implications for income, labour and health policies. This database represents a new and unique resource that will fill an important national data gap, and enable a wide range of relevant research.

*Corresponding Author:

Email Address: Claudia.sanmartin@canada.ca (C. Sanmartin)

