Use of the Population Grouping Methodology of the Canadian Institute for Health Information to predict high-cost health system users in Ontario

Sharada Weir DPhil, Mitch Steffler MA, Yin Li PhD, Shaun Shaikh MA, James G. Wright MD MPH, Jasmin Kantarevic PhD

ABSTRACT

BACKGROUND: Prior research has consistently shown that the heaviest users account for a disproportionate share of health care costs. As such, predicting high-cost users may be a precondition for cost containment. We evaluated the ability of a new health risk predictive modelling tool, which was developed by the Canadian Institute for Health Information (CIHI), to identify future high-cost cases.

METHODS: We ran the CIHI model using administrative health care data for Ontario (fiscal years 2014/15 and 2015/16) to predict the risk, for each individual in the study population, of

being a high-cost user 1 year in the future. We also estimated actual costs for the prediction period. We evaluated model performance for selected percentiles of cost based on the discrimination and calibration of the model.

RESULTS: A total of 11 684 427 individuals were included in the analysis. Overall, 10% of this population had annual costs exceeding $3050 per person in fiscal year 2016/17, accounting for 71.6% of total expenditures; 5% had costs above $6374 (58.2% of total expenditures); and 1% exceeded $22 995 (30.5% of total expenditures). Model performance increased with higher cost

thresholds. The c-statistic was 0.78 (reasonable), 0.81 (strong) and 0.86 (very strong) at the 10%, 5% and 1% cost thresholds, respectively.

INTERPRETATION: The CIHI Population Grouping Methodology was designed to predict the average user of health care services, yet performed adequately for predicting high-cost users. Although we recommend the development of a purpose-designed tool to improve model performance, the existing CIHI Population Grouping Methodology may be used — as is or in concert with additional information — for many applications requiring prediction of future high-cost users.

A substantial literature across health systems shows that the highest users of services account for disproportionate shares of the public costs of health care. It has recently been reported that more than three-quarters of individual health care costs in Ontario were incurred by just 10% of the population.[1] Similarly, an Ontario Ministry of Health and Long-Term Care (MOHLTC) analysis of inpatient and home care costs found that the top 5% of patients were responsible for 61% of spending in those domains.[2] Consistent findings have been reported for Manitoba, Alberta and British Columbia.[3–7]

Some of the highest-cost cases may be explained by rare, unpredictable events, but others arise in the presence of multiple chronic conditions. Research from the United States has suggested that spending on chronic conditions accounts for the

majority of health care expenditures.[8] Predicting high-cost users may help us to understand and better manage public spending on health care.

Cognizant of this need, the Ontario MOHLTC developed a predictive model for high-cost users based on sociodemographic, utilization and clinical diagnostic characteristics.[9] Although the model performed well, it relied on a coarse categorization of 20 diagnostic variables consisting of broadly defined chapters of the *International Statistical Classification of Diseases and Related Health Problems, 10th Revision* (ICD-10) and a small number of chronic conditions, which limited its utility for explaining predictions. Moreover, this model is not available for use outside the MOHLTC. As such, there is a need for a predictive model that can be applied more widely by

researchers and other stakeholders with an interest in health policy and spending in Canada.

The Canadian Institute for Health Information (CIHI) has recently released a new population-based case mix product, the Population Grouping Methodology, which uses diagnoses obtained from patient health care encounters in multiple settings to summarize the universe of diagnosis codes into a clinically meaningful set of 226 health conditions. The grouping and modelling methodologies are described in more detail in Appendix 1 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.191297/-/DC1) and in previous reports.[10,11] The CIHI grouping methodology was not designed to predict high-cost cases, and previous work has already shown that the model performs better for low- and moderate-cost users than for highest-cost users (i.e., those with annual costs exceeding $25 000).[11] We evaluated the suitability of CIHI's model for predicting future high-cost users in Ontario by examining the predicted costs for individuals who exceeded the top 10%, 5%, and 1% thresholds of actual cost.

Methods

Study design and setting

The study design involved validation of CIHI's cost-risk predictions for high-cost health system users in Ontario. The CIHI model was developed using a 3-province data set from 2010/11 through 2012/13. We used out-of-sample data representing the population of Ontario, one of the provinces that was included in model development. Using diagnoses recorded in the validation set over a look-back period of 2 fiscal years (2014/15 and 2015/16), we ran the CIHI model to predict high-cost individuals in the following year (2016/17) and then assessed model performance according to whether individuals who were predicted to have high costs actually had high costs, at selected cost thresholds (e.g., top 10%). To avoid including individuals who had died or left the province without their change in eligibility being recorded in the Registered Persons Database, we limited the study population to Ontario residents under 106 years of age who had health care utilization during the look-back (baseline) period and were eligible for coverage under the Ontario Health Insurance Plan (OHIP) during the full study period.

Data sources

The Ontario MOHLTC provided administrative health care eligibility and encounters files. We combined data from the Registered Persons Database for eligible individuals with all available encounters from CIHI's Discharge Abstract Database, CIHI's National Ambulatory Care Reporting System and the OHIP physician billings database. Together, these sources provide information on almost all inpatient, outpatient and ambulatory encounters in Ontario.

Identification of health conditions

We processed diagnosis data using CIHI's Population Grouping Methodology software, which output a set of 226 health condition flags (Appendix 2, Table A2-1, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.191297/-/DC1), along with resource intensity weights from a model predicting cost in the next year using health conditions found in the look-back period. Resource intensity weights are relative risk scores created by dividing each individual's predicted cost by the average for the population, so that individuals (and subgroups) can be compared in terms of their expected utilization of health care resources.

To understand better the drivers of cost and risk, we examined the prevalence of chronic conditions among patients in the top risk groups and among those with the highest actual cost. To identify chronic conditions, we included the 85 of the 226 health conditions that CIHI classified as being either chronic medical conditions, mental health conditions or cancer (Appendix 2, Table A2-2, contains details on the conditions and their classifications).

Estimation of costs

We estimated costs from the perspective of the public payer, the Ontario MOHLTC. Key components of what we denoted as "actual costs" in the prediction year were estimated according to commonly accepted costing methods[12] and were aligned with those used by CIHI. These values approximated those used by CIHI in model development and covered the same health care sectors (Appendix 3, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.191297/-/DC1).[10]

Statistical analysis

For analyses of model performance, we selected thresholds of actual cost representing high-cost users (top 10%, top 5% and top 1%), because users at these thresholds have been found in prior studies to account for the majority of health care expenditures.[1–7] Similarly, risk thresholds were defined for individuals in the top 10%, 5% and 1% in terms of predicted cost. At each threshold, we assessed the ability of the model to discriminate between higher and lower cost-risk individuals and examined how well the model was calibrated in terms of the difference between predicted and observed costs.

For model discrimination, we calculated the sensitivity (ability of the model to correctly select high-cost cases), specificity (ability of the model to accurately exclude low-cost individuals), positive predictive value (PPV; the proportion of cases predicted to be high cost that actually were high cost) and accuracy (correctly predicted cases — both high and lower cost — as a proportion of all individuals). We plotted the trade-off between sensitivity and specificity for each cost threshold using receiver operating characteristic curves. Strong model performance is indicated by values of the c-statistic, or area under the curve (AUC), of 0.8 or greater.[13] We also calculated the Brier score,[14] which measures the mean squared prediction error, where lower scores are better (within the range 0–1). We conducted sensitivity analyses by varying the population as follows: including individuals with no health care utilization in the concurrent (baseline) period or excluding individuals who died during the prospective period (i.e., those whose baseline data were collected within 12 months before their death).

Accuracy of the absolute estimates — or model calibration — was assessed by plotting predicted versus actual cost and visually examining the divergence at different levels of actual cost.

Ethics approval

Formal ethics approval was not required because this study used de-identified administrative health care data that were obtained from the Ontario MOHLTC under an agreement with the Ontario Medical Association.

Results

The Registered Persons Database included 13 293 352 people aged younger than 106 years as of the end of the concurrent period (Mar. 31, 2016) who were residents of Ontario during fiscal years 2014/15 and 2015/16, who had at least 1 encounter with the health care system in the prior 5 years and who were eligible for OHIP coverage in fiscal year 2016/17. Limiting the analyses to those with health care encounters during the baseline period removed 847 070 individuals (6.4% of the initial population). Removing late entrants and early exits (n = 761 855) to ensure sufficient data for prediction resulted in the exclusion of a further 5.7% of the initial population. The final study population consisted of 11 684 427 individuals who met all inclusion and exclusion criteria.

To highlight high-cost cases, we focused on the 90th, 95th and 99th percentiles of actual cost in the prospective period. Ten percent of the population (n = 1 168 441) were estimated to have costs in excess of $3050 per person in fiscal year 2016/17, accounting for 71.6% of total expenditures; 5% (n = 584 222) had costs greater than $6374, accounting for 58.2% of total expenditures; and 1% (n = 116 845) had costs exceeding $22 995, accounting for 30.5% of total expenditures.

Clinical profile of high-cost users

Table 1 provides descriptive information about the individuals predicted to have high costs compared with those who actually had high costs in the prospective period. The prevalence of all

Table 1: Patient demographic characteristics and clinical complexity (2-year prevalence of chronic conditions) by risk score and prospective cost thresholds

Characteristic	Risk score threshold*				Cost threshold†			
	≥ 99th percentile (top 1% by risk score)	≥ 90th percentile (top 10% by risk score)	≥ 50th percentile (top 50% by risk score)	< 50th percentile (bottom 50% by risk score)	≥ 99th percentile (top 1% by cost)	≥ 90th percentile (top 10% by cost)	≥ 50th percentile (top 50% by cost)	< 50th percentile (bottom 50% by cost)
Demographic								
Age, yr, mean ± SD	71.0 ± 16.7	63.8 ± 19.9	51.6 ± 21.6	32.5 ± 19.7	66.8 ± 18.4	56.2 ± 21.4	49.0 ± 22.4	35.2 ± 21.0
Sex, female, %	47.3	55.2	55.4	49.4	46.4	58.1	57.0	47.8
Clinical complexity, %								
≥ 1 chronic conditions	100.0	98.5	89.0	24.4	93.0	84.6	74.4	39.1
≥ 2 chronic conditions	99.9	89.8	56.9	2.5	79.4	63.5	46.8	12.7
≥ 3 chronic conditions	99.0	74.8	31.2	0.1	63.9	44.6	27.3	4.0
≥ 5 chronic conditions	84.7	36.7	8.9	0.0	36.1	19.2	8.4	0.5
≥ 10 chronic conditions	14.5	1.9	0.4	0.0	4.7	1.4	0.4	0.0
Severity,‡ % with any								
Minor chronic	90.4	78.4	66.3	17.4	75.1	65.7	56.4	27.5
Moderate chronic	77.5	52.6	28.3	1.7	49.0	36.7	23.6	6.4
Major chronic	92.7	52.9	17.6	0.0	52.2	27.1	14.9	2.7
Major mental health	27.7	17.8	5.0	0.0	12.6	7.6	4.3	0.8
Other mental health	38.7	30.7	26.1	6.1	27.2	27.3	22.3	9.9
Cancer (any type)	35.9	26.5	10.5	0.5	24.2	15.2	9.0	1.9

Note: SD = standard deviation.
*For risk score thresholds, the category "≥ 99th percentile" includes individuals whose risk scores were in the top 1% of the population; the category "≥ 90th percentile" includes individuals whose risk scores were in the top 10% of the population; the category "≥ 50th percentile" includes individuals whose risk scores were in the top half of the population; and the < 50th percentile includes individuals whose risk scores were in the bottom half of the population. Note that the category "≥ 50th percentile" includes individuals who were also in the 90th and 99th percentiles, and the category "≥ 90th percentile" includes individuals who were also in the 99th percentile.
†For cost thresholds, the category "≥ 99th percentile" includes individuals whose actual costs in the prospective year were in the top 1% of the population; the category "≥ 90th percentile" includes individuals whose actual costs in the prospective year were in the top 10% of the population; the "≥ 50th percentile" includes individuals whose actual costs in the prospective year were in the top half of the population; and the category "< 50th percentile" includes individuals whose actual costs in the prospective year were in the bottom half of the population. Note that the category "≥ 50th percentile" includes individuals who were also in the 90th and 99th percentiles, and the category "≥ 90th percentile" includes individuals who were also in the 99th percentile.
‡Of the 226 health conditions output by the Canadian Institute for Health Information Population Grouping Methodology, 85 were identified as chronic (including several mental health conditions and cancer, which was counted as a single condition regardless of the number of cancer diagnoses recorded in the data).

categories of chronic disease and multimorbidity (i.e., having 2 or more chronic conditions) rose starkly with predicted risk and actual cost. Just over one-tenth of those in the bottom half of the actual cost distribution had multiple comorbidities versus almost 80% of those in the top 1% cost group.

Model performance

Model discrimination measures for selected risk score cut points and cost thresholds are presented in Table 2. Model performance increased with higher cost thresholds, with AUC values of 0.78 (good performance), 0.81 (strong) and 0.86 (very strong) for the 90th, 95th and 99th percentile thresholds, respectively, indicating that cost thresholds became increasingly predictable in relation to prior diagnoses. Performance was somewhat affected by varying the population considered. The AUC value was slightly lower at the 95th percentile of cost when deaths were excluded from the analysis (AUC 0.79) compared with both the main analysis and the analysis that included non-users (AUC 0.81). Detailed results of the sensitivity analyses are presented in Appendix 4 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.191297/-/DC1).

Model calibration is illustrated in Figure 1. An ideal model would follow the dashed 45° line, whereby the observed

probability matches the predicted probability at every point. Because high-cost cases are rare events, most cases have low prediction probabilities (shown on the histograms in the bottom panel). The discrepancy between the logistic models and the ideal model increases with the prediction probability, yet the curves are all monotonically increasing, which means that the models are consistently predicting more high-cost cases with higher prediction probabilities. In each of the top 10%, the top 5% and the top 1% scenarios, both Brier and AUC scores showed a moderate-to-strong fit of the model. Nevertheless, performance at the high end of the charts — which is most relevant for policy purposes — was relatively poor compared with performance at the low end.

Interpretation

Heavy users of health care services have received a great deal of attention recently, as governments are concerned about the large and rising share of public expenditures devoted to health care. We found that the top 10% of high-cost cases accounted for almost three-quarters of total expenditures, and the top 1% subset accounted for nearly one-third. The accuracy of the CIHI model improved with increasing cost threshold. We found that chronic, serious conditions were common among those in the top percentiles of risk and cost, along with increased rates of multimorbidity, indicating that part of the utilization among the highest-cost users was predictable.

The Ontario MOHLTC previously created a high-cost user model (for internal use only), which included demographic variables (age, sex, rural v. urban residence), clinical variables (primarily ICD-10 chapters), socioeconomic status variables and health care utilization variables for the current and prior 2 years.[9] Performance of the CIHI model (without modifications) at the top 5% cost threshold, as reported here, was not quite as good as the MOHLTC's custom-built, high-cost case model. The c-statistic was 0.81 (strong), compared with 0.87 (very strong) for the MOHLTC model. Sensitivity at the top 5% risk threshold was 31.1% (v. 42.2%), specificity was 96.4% (v. 97.0%), positive predictive value was 31.1% (v. 42.6%), and model accuracy (true positives plus true negatives as a percentage of all predictions) was 93.1% (v. 94.2%). A purpose-designed high-cost model built upon the CIHI diagnosis grouping methodology platform with the addition of other predictors of high cost utilization (e.g., prior cost) may be expected to boost performance to levels comparable with, or exceeding, those reported elsewhere.[9,15]

Although the purpose-designed MOHLTC model had superior performance, health risk predictive modelling tools like CIHI's Population Grouping Methodology have been used in other jurisdictions to select cases for targeted care initiatives. For example, the Vermont Chronic Care Initiative reported savings net of program expenses in excess of US$30 million for fiscal year 2014/15,[16] which illustrates the potential value of such tools in supporting targeted programs to improve care coordination for residents eligible for state-funded services.[15] It is noteworthy that PPV — perhaps the most important metric in this

Table 2: CIHI Population Grouping Methodology model performance evaluated at the top 10%, 5% and 1% cost thresholds, by select risk score cut points			
	Risk score percentile; criterion value		
Threshold of actual cost in prospective period	**90th**	**95th**	**99th**
	Sensitivity*		
Top 10%	0.391	0.236	0.064
Top 5%	0.464	0.311	0.098
Top 1%	0.629	0.494	0.206
	Specificity†		
Top 10%	0.928	0.968	0.996
Top 5%	0.919	0.964	0.995
Top 1%	0.905	0.954	0.992
	Positive predictive value‡		
Top 10%	0.385	0.463	0.631
Top 5%	0.232	0.311	0.491
Top 1%	0.063	0.099	0.206
	Accuracy§		
Top 10%	0.872	0.892	0.899
Top 5%	0.896	0.931	0.950
Top 1%	0.903	0.950	0.984

Note: FN = false negatives, FP = false positives, TN = true negatives, TP = true positives.
*Sensitivity defined as proportion of high-cost cases predicted correctly (TP/[TP + FN]).
†Specificity defined as proportion of lower-cost cases predicted correctly (TN/[TN + FP]).
‡Positive predictive value defined as proportion of predicted high-cost cases that were high cost (TP/[TP + FP]).
§Accuracy defined as proportion of all cases predicted correctly, both high-cost and lower-cost ([TP + TN]/[TP + FP + TN + FN]).
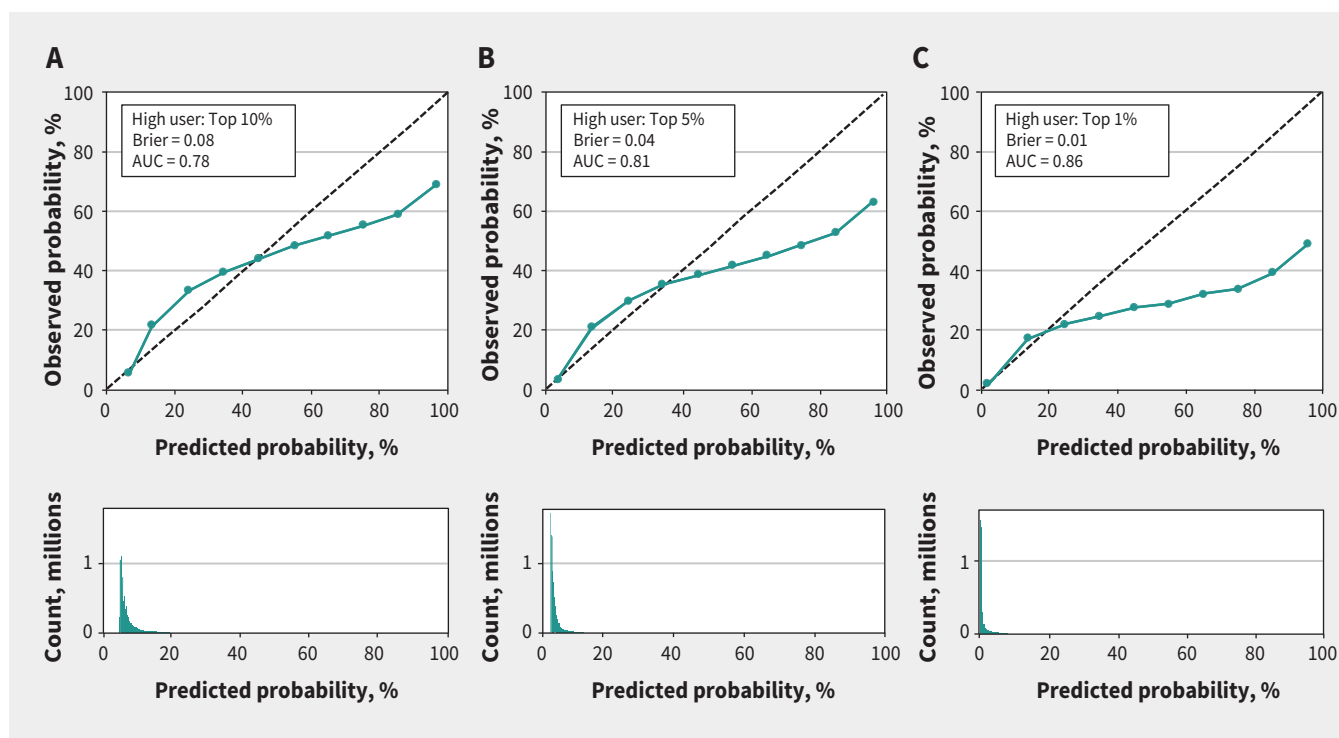
Figure 1: Model calibration, as predicted versus observed probabilities, for the top 10% (A), top 5% (B) and top 1% (C) cost thresholds. The Brier score (ranging from 0 to 1) measures the mean squared prediction error, where lower scores are better. For area under the receiver operating characteristic curve (AUC), values of 0.8 or greater indicate strong model performance. The population is mostly concentrated in the lower range of predicted cost, where the model for the top 1% performs best; as such, although this model is less well calibrated at the top end, it does the best overall in terms of the Brier score and AUC.

context, given that it indicates the percentage of accurately predicted high-cost users — for the CIHI grouping methodology at the 90th percentile cost threshold and 95th percentile of risk (PPV 0.463) was much higher than that reported in a prior study evaluating the case selection tool that was used in Vermont in 2012 (PPV 0.35; relevant model, Chronic Illness and Disability Payment System).[15] Like the conceptually equivalent CIHI tool, that model was not purpose-designed to predict the highest-cost cases, yet it served the purposes of the Vermont Chronic Care Initiative effectively.

The CIHI Population Grouping Methodology performed adequately for predicting high-cost users and may provide a helpful initial filter to select cases for interventions to enhance health care delivery to those in high need and potentially to contain future health care expenditures. Although we recommend the development of a purpose-designed tool to improve model performance, the existing CIHI Population Grouping Methodology may be used — as is or in concert with additional information — for many applications, such as case management and care coordination, that require prediction of future high-cost users.

Limitations

This study had several limitations. Ideally, costs would be estimated from a societal perspective and would include out-of-pocket expenditures by patients, formal and informal caregiving, and other relevant costs. However, data needed to include these costs were not available, both in the creation of the CIHI cost model and in our own analyses. In addition, neither the CIHI

model nor our analyses included costs associated with inpatient mental health stays, inpatient rehabilitation, home care or long-term care.

Our estimates of total health care costs were lower than those reported in another recent study for Ontario, for which data were available from all health care settings.[1] This was most apparent for the highest-cost users, who have been shown by prior research to be heavier-than-average users of non-acute hospital care, continuing care services and mental health care,[1,17] the settings that were not available for our study. The top 1% of high-cost users in the prior study had expenditures exceeding $44 906 in fiscal year 2011/12, which was more than double our result of $22 254 in fiscal year 2016/17, even without accounting for inflation. Nevertheless, the top 1% accounted for a similar percentage of expenditures (almost 31% in our data v. 33% in the prior study). Although CIHI's grouping methodology appears to be robust to incomplete data availability,[11] data limitations may have partially obscured the clinical profiles of some of the highest-cost users. Model performance might have been better if complete diagnostic information from all settings had been available for this study.

Conclusion

Although the CIHI model was optimized to maximize predictive performance for the average user of health care services, on the basis of only past diagnoses of health conditions, the model performed only slightly worse than the purpose-designed model used by the MOHLTC for predicting high-cost users. CIHI may

wish to consider inclusion of a purpose-designed model in future releases of its Population Grouping Methodology. Until then, the current model provides clinical condition information and risk scores that may provide a practical starting point for classification of patients according to cost risk.

References

1. Wodchis WP, Austin PC, Henry DA. A 3-year study of high-cost users of health care. *CMAJ* 2016;188:182-8.

2. Rais S, Nazerian A, Ardal S, et al. High-cost users of Ontario's healthcare services. *Healthc Policy* 2013;9:44-51.

3. Kozyrskyj A, Lix L, Dahl M, et al. *High-cost users of pharmaceuticals: Who are they?* Winnipeg: University of Manitoba, Faculty of Medicine, Department of Community Health Sciences, Manitoba Centre for Health Policy; 2005. Available: http://mchp-appserv.cpe.umanitoba.ca/reference/high-cost.pdf (accessed 2018 May 2).

4. Deber RB, Lam KCK. Handling the high spenders: implications of the distribution of health expenditures for financing health care. In: Chambers S, Jentleson BW, editors. *Proceedings of the 2009 American Political Science Association annual meeting*; 2009 Sept. 3–6; Toronto. Rochester (NY): Social Science Research Network; 2009. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450788 (accessed 2018 May 2).

5. Roos NP, Shapiro E, Tate R. Does a small minority of elderly account for a majority of health care expenditures? A sixteen year perspective. *Milbank Q* 1989;67:347-69.

6. Reid R, Evans R, Barer M, et al. Conspicuous consumption: characterizing high users of physician services in one Canadian province. *J Health Serv Res Policy* 2003;8:215-24.

7. Briggs T, Burd M, Fransoo R. Identifying high users of healthcare in British Columbia, Alberta and Manitoba. *Healthc Pap* 2014;14:31-6, discussion 58-60.

8. Conway P, Goodrich K, Machlin S, et al. Patient-centered care categorization of U.S. health care expenditures. *Health Serv Res* 2011;46:479-90.

9. Chechulin Y, Nazerian A, Rais S, et al. Predicting patients with high risk of becoming high-cost healthcare users in Ontario (Canada). *Healthc Policy* 2014;9:68-79.

10. Population grouping methodology [information sheet]. Ottawa: Canadian Institute for Health Information; 2017.

11. Li Y, Weir S, Steffler M, et al. Using diagnoses to estimate health care cost risk in Canada. *Med Care* 2019;57:875-81.

12. Wodchis WP, Bushmeneva K, Nikitovic M, et al. 4.8. Physician services. In: *Guidelines on person-level costing using administrative databases in Ontario*. Working Paper Series Vol. 1:35-6. Toronto: Health System Performance Research Network; 2013. Available: https://tspace.library.utoronto.ca/bitstream/1807/87373/1/Wodchis%20et%20al_2013_Guidelines%20on%20Person-Level%20Costing.pdf (accessed 2020 Apr. 2).

13. Hosmer DW, Lemeshow S. *Applied logistic regression*. 2nd ed. New York: John Wiley & Sons; 2000.

14. Rufibach K. Use of Brier score to assess binary predictions [letter]. *J Clin Epidemiol* 2010;63:938-9.

15. Weir S, Aweh G, Clark RE. Case selection for a Medicaid chronic care management program. *Health Care Financ Rev* 2008;30:61-74.

16. 7.6: Vermont Chronic Care Initiative (VCCI). In: *Vermont blueprint for health. 2015 annual report*. Waterbury (VT): Government of Vermont; 2016. p. 64-7. Available: https://ljfo.vermont.gov/assets/docs/healthcare/Health-Reform-Oversight-Committee/2016-Interim-Reports/da34dd8eeb/2015-Annual-Blueprint-Report.pdf (accessed 2020 Apr. 2).

17. Anderson M, Revie CW, Quail JM, et al. The effect of socio-demographic factors on mental health and addiction high-cost use: a retrospective, population-based study in Saskatchewan. *Can J Public Health* 2018;109:810-20.