

Using Diagnoses to Estimate Health Care Cost Risk in Canada

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

Objective: Until recently, the options for summarizing Canadian patient complexity were limited to health risk predictive modeling tools developed outside of Canada. This study aims to validate a new model created by the Canadian Institute for Health Information (CIHI) for Canada's health care environment.

Research Design: This was a cohort study.

Subjects: The rolling population eligible for coverage under Ontario's Universal Provincial Health Insurance Program in the fiscal years (FYs) 2006/2007–2016/2017 (12–13 million annually) comprised the subjects.

Measures: To evaluate model performance, we compared predicted cost risk at the individual level, on the basis of diagnosis history, with estimates of actual patient-level cost using “out-of-the-box” cost weights created by running the CIHI software “as is.” We next considered whether performance could be improved by recalibrating the model weights, censoring outliers, or adding prior cost.

Results: We were able to closely match model performance reported by CIHI for their 2010–2012 development sample (concurrent $R^2=48.0\%$; prospective $R^2=8.9\%$) and show that performance improved over time (concurrent $R^2=51.9\%$; prospective $R^2=9.7\%$ in 2014–2016). Recalibrating the model did not substantively affect prospective period performance, even with the addition of prior cost

and censoring of cost outliers. However, censoring substantively improved concurrent period explanatory power (from $R^2=53.6\%$ to 66.7%).

Conclusions: We validated the CIHI model for 2 periods, FYs 2010/2011–2012/2013 and FYs 2014/2015–2016/2017. Out-of-the-box model performance for Ontario was as good as that reported by CIHI for the development sample based on 3-province data (British Columbia, Alberta, and Ontario). We found that performance was robust to variations in model specification, data sources, and time.

Key Words: CIHI Population Grouping Methodology, risk adjustment, model validation

(*Med Care* 2019;57: 875–881)

Health risk predictive models summarize clinical complexity at the patient level. Methods vary, but a key approach involves regressing total health care costs as a function of age and sex plus a set of health conditions derived from diagnoses recorded on health care claims. Cost predictions derived from the regression coefficients are divided by the average cost to create relative risk scores for each member of the population. Such models were developed initially in the 1980s and 1990s in the United States in an effort to improve upon the simple demographic models used at the time to adjust capitated payments.^{1,2} In the intervening years, US models have been adopted by Canadian researchers eager to take advantage of the strengths of clinical grouping methodologies to summarize patient case-mix.^{3–5} However, differences in health care delivery between the United States and Canada call into question the suitability of applying US models in a Canadian context.

The Canadian Institute for Health Information (CIHI) has produced the first-ever Canadian health risk modeling suite.⁶ CIHI's Population Grouping Methodology was developed in consultation with a team of clinical experts to address the need for a set of predictive models that reflect the Canadian health care landscape (Appendix, Supplemental Digital Content 1, <http://links.lww.com/MLR/B871>).^{2,7,8} As CIHI used pooled data from 3 provinces—each with their own unique characteristics—and the model may in the future be used to risk-adjust capitated payments to physicians, there was a need to evaluate model performance specifically in the Ontario context.

This study aimed to validate the use of CIHI's case-mix methods for the purposes of predicting cost in Ontario. Our key research questions are whether the CIHI model performance for Ontario alone matches that reported previously by CIHI for the

From the Department of Economics, Policy & Research, Ontario Medical Association, Toronto, ON, Canada.

Y.L. and S.W. are co-first authors.

The Population Grouping Methodology is owned by the Canadian Institute for Health Information (CIHI) and used under license. Population-level health care data were obtained from the Ontario Ministry of Health and Long-Term Care (MOHLTC) under an agreement with the Ontario Medical Association (OMA). Both CIHI and the MOHLTC had no involvement in or control over the design and conduct of the study; the collection, analysis, and interpretation of the data; the preparation of the data; the decision to publish; or the preparation, review, and approval of the manuscript.

The authors declare no conflict of interest.

Reprints: Sharada Weir, DPhil, Department of Economics, Policy & Research, Ontario Medical Association, 150 Bloor Street West, Suite 900, Toronto, ON, Canada M5S 3C1. E-mail: sharada.weir@oma.org.

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website, www.lww-medicalcare.com.

Copyright © 2019 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

ISSN: 0025-7079/19/5711-0875

3-province validation data set, whether model performance is durable over time, and whether it is possible to improve performance by recalibrating the model using Ontario data alone or by adding prior cost to the model. This represents the first published account of CIHI's health risk predictive model performance. We hope that this effort will increase confidence in adopting the methodology among public health agencies, government, and other stakeholders seeking to improve the set of planning tools available to them.

METHODS

Patient risk scores were assigned by the CIHI grouper using diagnostic information drawn from several sources. The Registered Persons Database (RPDB) was used to assess eligibility for inclusion in the study population and assign patients to age-sex groups. The study population was linked to the Ontario health administrative data at the individual level using unique encrypted identifiers. Key files included the following (and they are): CIHI's Discharge Abstract Database (DAD), which contains diagnostic records from hospital discharges in Ontario; CIHI's National Ambulatory Care Reporting System (NACRS), which includes most hospital-based and community-based ambulatory care records (ie, day surgery, outpatient and community-based clinics, and emergency departments; and the Ontario Health Insurance Plan (OHIP) claims database, which is made up of physician billing records across settings.

The study population included all individuals living in Ontario who were eligible and registered for coverage under OHIP during the study period. We restricted the population to those aged ≤ 105 years to limit the likelihood of including deceased individuals whose deaths were not recorded in our database. Similarly, people who had no contact with the health care system within the past 5 years were excluded from the cohort for a given year, as they may have either left the province or died without their status change being recorded in our data.

The model was validated for 2 periods: fiscal years (FYs) 2010–2012 and 2014–2016. The first corresponded to the model development period and the second was included to determine whether the model remained valid over time. Diagnoses captured during the concurrent periods (FYs 2010/2011–2011/2012 and FYs 2014/2015–2015/2016) were used to explain costs that were incurred in the same periods and to predict costs for the corresponding prospective periods (FYs 2012/2013 and 2016/2017).

The CIHI model was run to produce individual risk scores. These values were then divided by the mean risk from the population to create normalized risk scores, which re-centered the average score for the population to 1.0. Thus, an individual with a risk score of 2.0 would have twice the cost risk compared with the average, whereas an individual with a risk score of 0.5 would have half the average risk.

We estimated actual costs as the sum of costs for acute inpatient hospitalizations, day surgery, hospital clinic visits, emergency department visits, and physician services using standard costing methods.⁹ For inpatient, day surgery, outpatient, and emergency department encounters, costs were calculated by multiplying the appropriate resource intensity weight for each

record in the Discharge Abstract Database (DAD) or National Ambulatory Care Reporting System (NACRS) by the Cost of a Standard Hospital Stay (CIHI 2018). Physician services were costed using the "fee paid" field attached to each fee-for-service physician services record or the "fee approved" field for shadow billed claims. Actual costs were then divided by the average cost for the population to produce a mean of 1.0, which is on the same scale as the predicted risk scores.

To evaluate model performance, we compared the predicted cost at the individual level derived from model risk scores with estimates of actual patient-level cost for the period. Model performance was evaluated using the coefficient of multiple determination (R^2), which is calculated as the square of the coefficient of multiple correlations, and the mean absolute prediction error, a measure which is less sensitive to outliers, calculated as the average of the absolute value of the prediction errors.

Several approaches to evaluating the CIHI Population Grouping Methodology were undertaken to explore different aspects of model performance. We first evaluated the "out-of-the-box" cost weights created by running the CIHI model grouper software "as is." To illustrate the gains obtained by including health conditions to predict cost, compared with predictions based on demographic data alone, we created age-sex only models for the concurrent and prospective periods (using 21 age groups for each sex).

We next considered whether performance could be improved by recalibrating the model weights using Ontario data alone. Recalibration was achieved by splitting the population for each concurrent period into development sets and validation sets. Three development sets were created using stratified random sampling to select 70% of: (1) health system users with at least 1 health condition; (2) users with no conditions; and (3) nonusers. Total actual costs in the development data were modeled separately for each of these groups, reproducing CIHI's methods. In keeping with CIHI's approach, some model weights were restricted to avoid negative cost predictions.¹⁰ The final recalibrated model weights were applied to the validation data set to create risk scores.

The second set of recalibrated prospective models was estimated by adding the prior-year cost, which has been found to be highly predictive of future cost in previous research.^{11,12} We also examined the influence of outliers by re-estimating the models with costs censored at CA\$100,000.

Finally, we explored the impact of the lookback period on model performance. Whereas the standard CIHI model uses 2 years of data to train their model, we ran the out-of-the-box model using diagnosis codes obtained from looking back over the prior 1, 2, 3, 4, or 5 years to determine whether there are returns to additional diagnostic information in explaining concurrent period costs and predicting prospective period costs. This exercise was repeated using the base FYs 2010 (anchor date: March 31, 2011) through 2015 (anchor date: March 31, 2016) to determine whether the impact of the lookback period choice exhibited a trend or was stable over time.

All analyses were carried out using SAS software, version 9.4 for Windows (SAS Institute, Cary, NC) and Tableau Desktop Professional Edition, version 10.5.5 (Tableau Software, Seattle, WA).

TABLE 1. Study Population

	Study Period 1 (Development Sample Period)		Study Period 2 (Most Recent Available Data Period)	
	Concurrent (FYs 2010/2011 and 2011/2012)	Prospective (FY 2012/2013)	Concurrent (FYs 2014/2015 and 2015/2016)	Prospective (FY 2016/2017)
No. observations	12,826,542	12,547,258	13,293,352	12,991,652
Age (mean) (y)	40.0	39.6	41.0	40.6
Age (SD) (y)	23.0	22.8	23.4	23.1
Aged 65+ (%)	15.90	15.12	17.73	16.94
Sex, male (%)	48.69	48.67	48.72	48.68
Health system users—with conditions (%)	89.15	89.20	88.98	89.02
Health system users—without conditions (%)	4.86	4.87	4.65	4.65
Health system nonusers (%)	5.98	5.93	6.37	6.33
Concurrent risk score (mean)	1.138		1.204	
Prospective risk score (mean)		1.045		1.098

FY indicates fiscal year.

RESULTS

Basic descriptive information on the study populations in each period is provided in Table 1. The population grew while becoming slightly older, on average, over the 2 study periods, but health system users comprised a slightly smaller share of the population in the later period. At the same time, risk scores rose. That is, the predicted costs of the CIHI model increased over time, reflecting greater clinical complexity of the population, on average, in the latter period.

Out-of-the-box model performance is presented in Table 2. The mean absolute prediction error column shows that costs were underestimated in the concurrent period and slightly overestimated in the prospective period, on average. The Ontario R^2 values corresponding to CIHI's development period were almost identical to those reported by CIHI for the full 3-province development sample.⁶ A few years later, Ontario R^2 values were slightly higher, showing that the model maintained predictive power over time.

Table 3 illustrates the range of model performance possibilities using recent data, starting from a simple demographic model and moving to a set of models calibrated on Ontario data and optimized using censoring of outliers and, for the prospective prediction models, the inclusion of the prior-year costs. The age-sex only model performed quite poorly compared with the out-of-the-box diagnostic model in both the prospective period ($R^2 = 0.4\%$ vs. 9.7%) and the concurrent period ($R^2 = 1.0\%$ vs. 52.7%). However, recalibrating the CIHI model using Ontario-only data for the most recent period did not substantively affect our results. Although censoring improved explanatory power in the concurrent period,

neither including prior cost nor censoring data to remove the influence of outliers had a substantial impact on model performance in the prospective period.

Next, we considered how well the model predicted users with different levels of predicted cost. Figures 1A and B illustrate the out-of-the-box concurrent and prospective model performance, respectively, by predicted cost category. Costs were more accurately predicted for low-cost users than for higher-cost users and were particularly underpredicted in the prospective period for the highest-cost users.

Figures 2A and B present the effect on R^2 of varying the concurrent period from 1 to 5 years. There were 2 key findings. First, model performance generally improved over time. Second, for the most recent 4 years, concurrent period performance improved with each additional year of diagnostic lookback. However, results before 2013 were less stable. The prospective prediction was maximized using 2 years of lookback and progressively deteriorated with additional years of diagnostic data. However, using only 1 year of lookback resulted in consistently lower R^2 values than with 2 years.

DISCUSSION

To our knowledge, no prior studies have been published independently evaluating the CIHI population grouper/health risk predictive model performance. CIHI reports that their model explained nearly half of the variance in cost using health conditions and condition interactions in the concurrent period ($R^2 = 47.5\%$) and almost one tenth in the

TABLE 2. CIHI Out-of-the-Box Model Validation Using Ontario-Data Only: Development Sample Period (FYs 2010/2011–2012/2013) Versus a More Recent Period (FYs 2014/2015–2016/2017)

Model and Period	No. Patients	Average Total Cost (\$)	Mean Absolute Prediction Error (%)	Coefficient of Multiple Determination (R^2) (%)
Concurrent, FYs 2010/2011–2011/2012	12,826,542	2798	60.21	48.01
Prospective, FY 2012/2013	12,826,542	1429	111.41	8.89
Concurrent, FYs 2014/2015–2015/2016	13,293,352	3044	59.05	51.88
Prospective, FY 2016/2017	13,293,352	1495	109.49	9.70

CIHI indicates Canadian Institute for Health Information; FY, fiscal year.

TABLE 3. Model Performance For Out-of-the-Box Versus Recalibrated Cost Weights, Validation Sample, FYs 2014/2015–2016/2017

Model	Concurrent (FYs 2014/2015– 2015/2016) (%)		Prospective (FY 2016/2017) (%)	
	MAPE	R ²	MAPE	R ²
Out-of-the-box model*	59.01	52.71	109.45	9.74
Age-sex only model	116.32	0.97	126.33	0.42
Recalibrated model, without prior cost	57.88	53.64	108.12	9.61
Recalibrated model, with prior cost			107.65	9.96
Recalibrated model, without prior cost, censored at CA \$100,000	53.54	66.74	108.25	8.91
Recalibrated model, with prior cost, censored at CA\$100,000			107.54	9.28

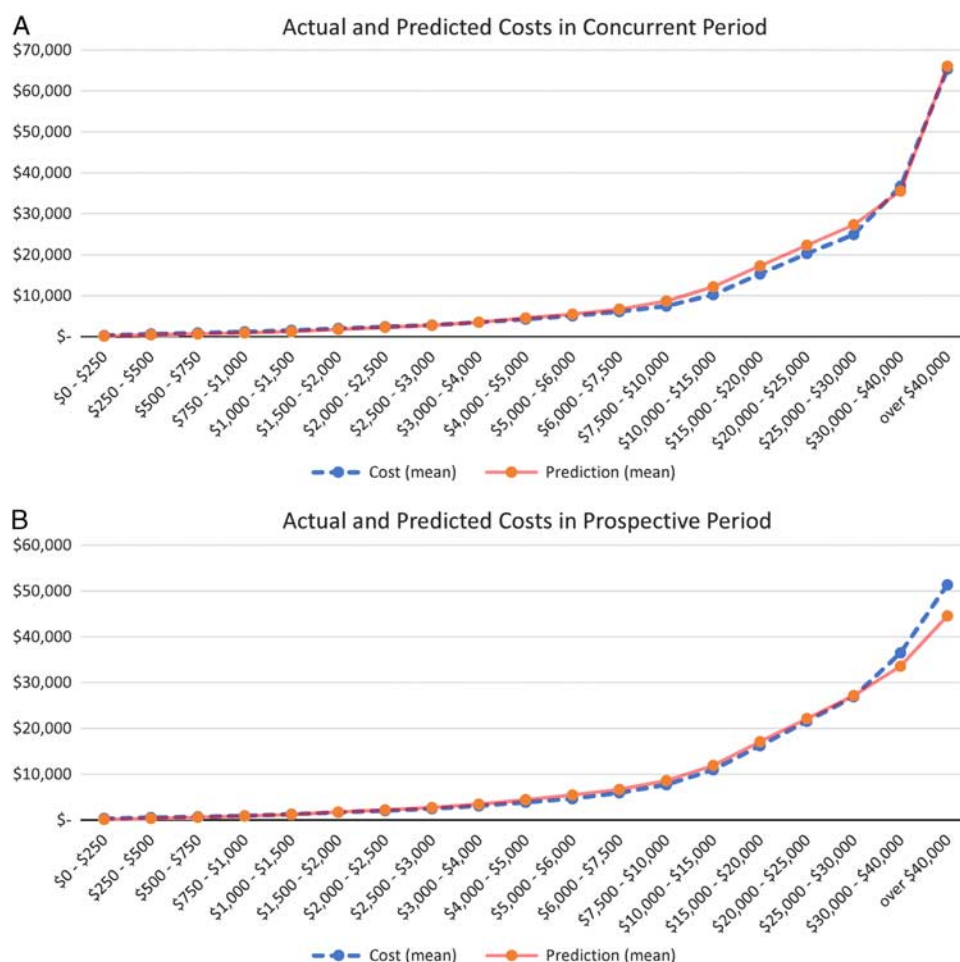
*Performance is reported for validation sample only, which differs slightly from the performance reported in Table 2 for the entire population.

FY indicates fiscal year; MAPE, mean absolute prediction error.

prospective period ($R^2=9.4\%$) for their 3-province model validation sample.⁶ We were able to closely match these results using data for Ontario alone for both the development sample period (FYs 2010/2011–2012/2013) and a more recent, out-of-development-sample period (FYs 2014/2015–2016/2017).

One question that may become particularly relevant in Ontario is whether the CIHI model performance in predicting future cost using health care diagnoses alone is adequate to permit its use in risk-adjusting prospective payments to physicians participating in capitated payment plans. It is instructive to compare the performance of the CIHI model with US performance of DxCG's Diagnostic Cost Group Hierarchical Condition Category (DCG-HCC) model, as this is the methodology most similar to that used by CIHI and is the one used to risk-adjust payments to Medicare Advantage capitated health plans in the United States.

The CIHI concurrent model results are nearly as good as those reported recently for the DxCG model ($R^2=52.6\%$), but the prospective model falls short of DxCG ($R^2=18.6\%$).¹³ CIHI's prospective model results are compared with earlier

**FIGURE 1.** Concurrent (A) and prospective (B). Actual versus predicted costs by level of predicted cost.

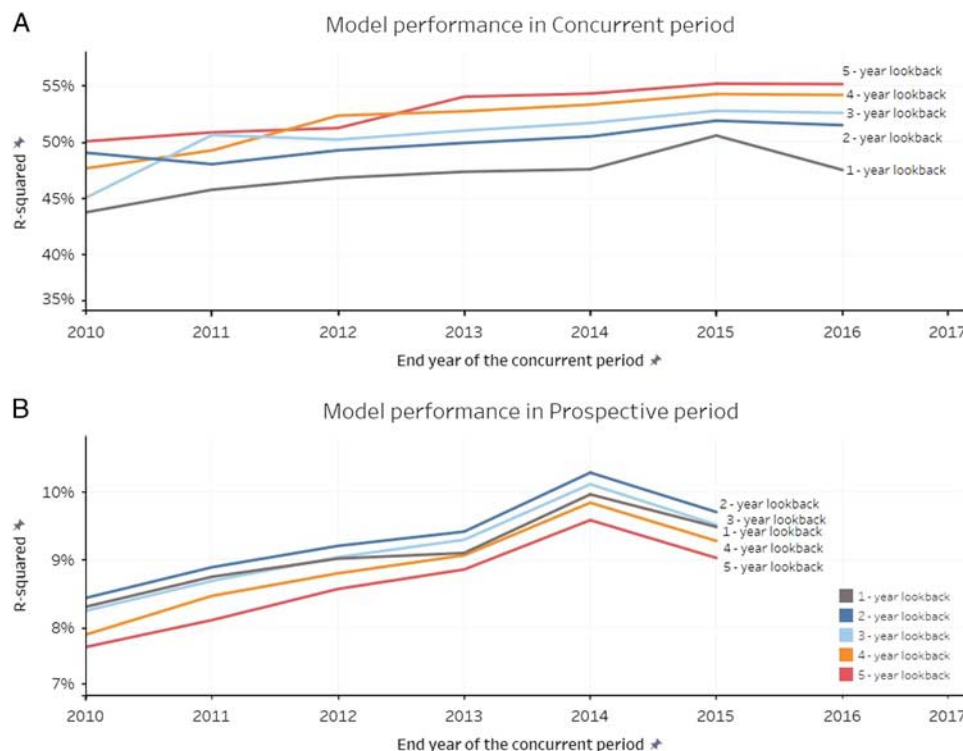


FIGURE 2. Concurrent (A) and prospective (B). Variation in model performance (R^2) by base year and lookback duration.

versions of the DxCG model (R^2 range: 8.1%–11.2%),^{1,2,14} before its adoption for risk-adjusting payments in the United States. In out-of-sample commercial health plan data validation of the DxCG methodology, the Society of Actuaries reported that the prospective payment model R^2 rose from 14.3% for the calendar years 1998–1999¹⁵ to 20.6% in the calendar years 2003–2004¹¹ and fell slightly to 18.6% in the calendar years 2012–2013.¹³ Improvements occurred following the adoption and subsequent expansion of DxCG methodology to risk-adjust payments. To put this performance in context, Newhouse et al¹⁶ estimated the maximum possible R^2 that may be achieved in a health condition prospective model between 20% and 30%.

A study sponsored by the Society of Actuaries suggested that significant gains in the predictive power of diagnosis-based risk adjustment models may be explained by a combination of improvements in data reporting and model refinements.¹⁴ However, the original model developers showed that changes to the DxCG grouper methodology between 1998 and 2000 had a negligible effect on the prospective R^2 , whereas running the same model with more recent data did improve performance. The authors speculated that improvements in diagnostic coding over time may help to explain gains in predictive power.²

Several studies have shown that, following the adoption of risk-adjusted payments to capitated health insurance plans by the US Centers for Medicare and Medicaid Services (CMS), diagnostic coding intensified among physicians in capitated plans.^{17–19} To the extent that changes in coding practices reflected more accurate coding (rather than fraudulent upcoding of diagnoses), model

performance would have improved, which may explain the higher R^2 in later years.

If so, and if payment reform is enacted in Ontario so that prospective payments in capitated primary care models become partly reliant on clinical complexity, we may expect CIHI's prospective model performance to improve over time.

However, it is also possible that the lower R^2 for the CIHI model may partly reflect greater variability in cost for a full population model in Canada versus models segregated by population characteristics of the insurance segment in the United States. The US Medicare, Medicaid, and private insurance populations were modeled separately and are more homogenous than the universally insured provincial populations in Canada. If that is the case, expectations with regard to the future performance of Canadian prospective payment models should be moderated.

Demographic-based Versus Diagnoses-based Risk Adjustment

One difference between the CIHI model and methodologically similar US models is that the US models include age and sex variables in addition to health conditions and health condition interactions. To investigate the potential role of demographic characteristics in predicting costs, we created a comparison model that was based on age and sex alone. We found that the demographic model predicted <1% of the variation. Moreover, unlike in the US models,¹⁰ which contained a smaller number of health conditions and health condition interactions than the CIHI model, we found that age and sex group coefficients were not significant in recalibrated health condition models in Ontario.

The weak performance of the demographic model is a concern for the Ontario health system, and one that echoes prior research showing that neither age and sex nor socioeconomic and mortality data were adequate for needs-based capitation.²⁰ Currently, physicians participating in capitated payment plans receive prospective payments based solely on the age-sex composition of their patient panels. Physicians who enroll patients who are more clinically complex, on average, than is typical will be underpaid in this framework. Hence, there are disincentives for physicians in such plans to take on too many patients with multiple chronic conditions. They are more likely to remain in fee-for-service, which some authors maintain offers no motivation to increase efficiency in treating the whole patient.^{21,22}

Model Recalibration

Although it is not appropriate to include prior cost in a payment model, as doing so may encourage inefficient resource use, models that include these data may be used in a variety of applications, including predicting high-cost cases for disease management programs. Research from the United States has suggested that better performance may be achieved by trimming costs to reduce the influence of cost outliers or by including prior-year costs in the model. The concurrent period performance was improved by censoring outliers. However, we did not find this for the prospective period model using Ontario data. Model performance was virtually unchanged by inclusion of prior cost or censoring high-cost cases. These findings enhance our confidence in the robustness of the out-of-the-box cost models produced by CIHI.

Optimal Lookback Period For Prediction

We found that the optimal lookback period varied for the concurrent versus prospective models. In the case of the concurrent period model, in which diagnoses recorded during the lookback period explain costs during the same period, the inclusion of up to 5 years of data improved model performance, as this ensured that diagnoses of chronic conditions that might not be recorded at every physician encounter were included. However, for prediction of costs 1 year into the future, in the prospective model, having a longer lookback period did not improve model performance. It seems that recently-diagnosed acute and chronic conditions are more predictive of future costs than are more distantly diagnosed conditions. In practice, the choice of the lookback period often is limited by data availability. Fortunately, in the context of the provincial health insurance model found in Canada, most patients maintain coverage and may be followed-up for multiple years. In this context, we recommend obtaining 2 years of lookback data at minimum for both models. However, if more data are available, these may be used to improve performance of the concurrent model.

Limitations of the Study Design, Data Sources, and Analytic Methods

An advantage of using CIHI's Population Grouping Methodology was that all diagnoses codes available from all available encounters were considered in calculating overall health risk. However, the utility of the model output depends upon the completeness and accuracy of the diagnosis codes

that are input into the model. Disease classification using administrative records is subject to additional sources of error. For example, a physician may suspect a particular diagnosis based on symptoms that the patient presents with and may record that diagnosis before it has been confirmed. The CIHI model attempts to validate physician coding by applying "tagging rules," which include ensuring that a minimum number of instances of a particular diagnosis are recorded in separate physician encounters and checking that the diagnosis is clinically feasible given the age and sex of the patient. However, these rules may have created false negatives while not eliminating all false-positive diagnoses in the data. These issues may have constrained model performance during the study period.

Our findings may also have been affected by the lack of data from mental health hospitalizations (Ontario Mental Health Reporting System), which was not available to the authors at the time of the study. However, given that we were able to nearly replicate the model performance statistics reported by CIHI for the model development set, missing data on inpatient mental health diagnoses do not seem to have substantively affected our results. This may be a testament to the robustness of the methodology.

These analyses take the perspective of the public payer, the Ontario Ministry of Health and Long-Term Care. Utilization and expenditure on out-of-pocket costs borne by patients, and any costs for segments of the population not covered by the Ministry, were not considered here, as these data were not readily available and were not used by CIHI to develop the Population Grouping Methodology.

Implications

To our knowledge, this study represents the first effort to publish validation results on the use of the CIHI Grouping Methodology to summarize clinical cost risk. The CIHI model suite has numerous potential applications for researchers and policymakers in Canada wishing to better understand patient case-mix and health system costs. Risk scores produced by the model may be used to adjust analyses of costs or outcomes of health care to account for average patient complexity, to select complexity-matched control subjects for research studies, to adjust payments to physicians (or other agents) in capitated payment plans, among other uses. Our validation of the model for a single province across different time periods serves to illustrate the robustness of the condition grouping and risk scoring methodology.

REFERENCES

1. Ash AS, Byrne-Logan S. How well do models work? Predicting health care costs. *Proceedings of the Section on Statistics in Epidemiology* (American Statistical Association); 1998:42-49.
2. Pope GC, Ellis RP, Ash AS, et al. Diagnostic cost group hierarchical condition category models for medicare risk adjustment: final report. Prepared for Health Care Financing Administration; 2000.
3. Austin PC, van Walraven C, Wodchis WP, et al. Using the Johns Hopkins Aggregated Diagnosis Groups (ADGs) to predict mortality in a general adult population cohort in Ontario, Canada. *Med Care*. 2011;49: 932-939.
4. Sibley LM, Moineddin R, Agha MM, et al. Risk adjustment using administrative data-based and survey-derived methods for explaining physician utilization. *Med Care*. 2010;48:175-182.

5. Reid RJ, MacWilliam L, Verhulst L, et al. Performance of the ACG case-mix system in two Canadian provinces. *Med Care*. 2001;39: 86–99.
6. Canadian Institute for Health Information (CIHI). CIHI's population grouping methodology 1.1 (compiled code): methodology report; 2017.
7. Buchner F, Goepffarth D, Wasem J. The new risk adjustment formula in Germany: implementation and first experiences. *Health Policy*. 2013;109: 253–262.
8. Cheng S, Austin P, Wodchis W, et al. Evaluation of population groupers. ICES Report; 2016.
9. Wodchis WP, Bushmeneva K, Nikitovic M, et al. *Guidelines on Person-Level Costing Using Administrative Databases in Ontario Working Paper Series Vol 1*. Toronto, ON: Health System Performance Research Network; 2013.
10. Ash AS, Ellis RP, Pope GC, et al. Using diagnoses to describe populations and predict costs. *Health Care Financ R*. 2000;21:7–28.
11. Winkelman R, Mehmud S. A comparative analysis of claims-based tools for health risk assessment. Society of Actuaries Sponsored Research Project; 2007. Available at: www.soa.org/research-reports/2007/hlth-risk-assessment/. Accessed April 23, 2018.
12. Chechulin Y, Nazerian A, Rais S, et al. Predicting patients with high risk of becoming high-cost healthcare users in Ontario (Canada). *Health Policy*. 2014;9:68–79.
13. Hileman G, Steele S. Accuracy of claims-based risk scoring models. Society of Actuaries Sponsored Research Project; 2016. Available at: www.soa.org/research-reports/2016/2016-accuracy-claims-based-risk-scoring-models/. Accessed April 23, 2018.
14. Ellis RP, Pope GC, Iezzoni LI, et al. Diagnosis-based risk adjustment for Medicare capitation payments. *Health Care Financ Rev*. 1996;17:101–128.
15. Cumming R, Knutson D, Cameron B, et al. A comparative analysis of claims-based methods of health risk assessment for commercial populations. Society of Actuaries; 2002. Available at: www.soa.org/Files/Research/Projects/2005-comp-analysis-methods-commercial-populations.pdf. Accessed April 23, 2018.
16. Newhouse JP, Manning WG, Keeler EB, et al. Adjusting capitation rates using objective health measures and prior utilization. *Health Care Financ Rev*. 1989;10:41–54.
17. Burns A, Hayford T. *Effects of Medicare Advantage Enrolment on Beneficiary Risk Scores*. Washington, DC: Congressional Budget Office Working Paper Series 2017-08; 2017.
18. Geruso M, Layton T. *Upcoding: Evidence From Medicare on Squishy Risk Adjustment*. Cambridge MA: National Bureau of Economic Research Working Paper 21222; 2018.
19. Kronick R, Welch WP. Measuring coding intensity in the Medicare Advantage Program. *Medicare Medicaid Res Rev*. 2014;4:a06.
20. Hutchison B, Hurley J, Birch S, et al. Needs-based primary medical care capitation: development and evaluation of alternative approaches. *Health Care Manag Sci*. 2000;3:89–99.
21. Martin KE, Rogal DL, Arnold SB. Health-based risk assessment: risk-adjusted payments and beyond. changes in health care financing & organization report. Academy Health; 2004.
22. Sevcik AE, Abu-Jaber T, Marek L. Understanding approaches to case-mix assessment and case-mix adjustment. Healthcare Quality Research Toolbox; JHQ Online 2004. W5-24-W5-29.