RESEARCH PAPER

Predicting Patients with High Risk of Becoming High-Cost Healthcare Users in Ontario (Canada)

Détecter les patients qui présentent un haut risque de devenir des usagers très coûteux pour les services de santé en Ontario (Canada)

YURIY CHECHULIN, MPH, MD
Senior Methodologist, Health Analytics Branch
Ontario Ministry of Health and Long-Term Care
Toronto, ON

AMIR NAZERIAN, MSC
Methodologist, Health Analytics Branch
Ontario Ministry of Health and Long-Term Care
Toronto, ON

SAAD RAIS, MSC
Senior Methodologist, Health Analytics Branch
Ontario Ministry of Health and Long-Term Care
Toronto, ON

KAMIL MALIKOV, CMA, MBA, MD
Senior Manager, Health Analytics Branch
Health System Information Management and Investment
Ontario Ministry of Health and Long-Term Care
Toronto, ON

Abstract

Literature and original analysis of healthcare costs have shown that a small proportion of patients consume the majority of healthcare resources. A proactive approach is to target interventions towards those patients who are at risk of becoming high-cost users (HCUs). This approach requires identifying high-risk patients accurately before substantial avoidable costs have been incurred and health status has deteriorated further. We developed a predictive

model to identify patients at risk of becoming HCUs in Ontario. HCUs were defined as the top 5% of patients incurring the highest costs. Information was collected on various demographic and utilization characteristics. The modelling technique used was logistic regression. If the top 5% of patients at risk of becoming HCUs are followed, the sensitivity is 42.2% and specificity is 97%. Alternatives for implementation of the model include collaboration between different levels of healthcare services for personalized healthcare interventions and interventions addressing needs of patient cohorts with high-cost conditions.

Résumé

La littérature et l'analyse des coûts des services de santé démontrent qu'une petite portion de patients mobilise la majorité des ressources des services de santé. Une démarche proactive consiste à privilégier les interventions visant des patients qui présentent des risques de devenir des utilisateurs très coûteux (UTC). Cette démarche demande un recensement précis des patients à haut risque avant que des coûts substantiels ne soient engagés et que leur état de santé ne se soit détérioré davantage. Nous avons mis au point un modèle de prévision pour recenser les patients susceptibles de devenir des UTC en Ontario. Les UTC correspondent aux premiers 5 % d'utilisateurs qui génèrent les coûts les plus élevés. L'information a été recueillie selon diverses caractéristiques démographiques et modes d'utilisation. La régression logistique a été employée comme technique de modélisation. Si on effectue le suivi des premiers 5 % de patients à risque de devenir des UTC, le résultat de la sensibilité est de 42,2 % et celui de la spécificité s'élève à 97 %. Les choix pour l'application du modèle comprennent la collaboration entre divers niveaux de services de santé pour offrir des interventions personnalisées, ou encore la mise en place d'interventions qui répondent aux besoins de groupes de patients présentant des états de santé dont les coûts sont élevés.

Literature from different jurisdictions has shown that a relatively small proportion of patients consume the majority of healthcare resources. In Ontario, for example, 5% of healthcare users consumed 61% of hospital and home care spending (Rais et al. 2013). A study of physician services utilization in British Columbia found that 5% of healthcare users consumed 30% of spending on physician services (Reid et al. 2003). In Manitoba, 5% of prescription drug users accounted for 41% of prescription expenditures (Kozyrskyj et al. 2005). In Australia, high-cost users (HCUs) accounted for 38% of both in-patient costs and in-patient days (Calver et al. 2006). US data from the Arizona Health Care Cost Containment System showed that 10% of patients accounted for two-thirds of healthcare costs (Moturu et al. 2010). Another study in the United States looking at healthcare expenditures from 1928 through 1996 found that the top 5% of HCUs accounted for more than half of health spending in both 1987 and 1996, while the top 10% accounted for about 70% of all healthcare spending (Berk and Monheit 2001). Yet another US study found that

5% of the population accounted for 49% of total healthcare spending (Center for Healthcare Research and Transformation 2010).

Given the impact of these HCUs of healthcare, interventions directed at them may improve overall patient outcomes and quality of life, and reduce healthcare spending. A report from The Commonwealth Fund (2012) supports this view, emphasizing the need to address high-cost healthcare users as the first step towards achieving "rapid improvements in the value of services provided." Gawande (2011) also argued that focusing on a few areas or individuals would have a significant impact on patient outcomes and system costs.

A proactive approach to addressing the problem of HCUs is to target interventions towards patients who are at risk of becoming HCUs. This approach is aimed at preventing at-risk patients from becoming HCUs in the first place. Such an approach requires some mechanism to identify or predict high-risk patients accurately before substantial preventable or avoidable costs have been incurred and health status has deteriorated further (Billings et al. 2006). One such mechanism is a statistical predictive model.

A number of publications have proposed various methods to predict future HCUs, each advocating a different model, predictor variables and type of data. Billings and colleagues (2006) presented a case-finding tool for patients at risk of readmission to hospital and developed an algorithm to identify high-risk patients in the United Kingdom. The key factors predicting subsequent admission included age, sex, ethnicity, number of previous admissions and clinical condition. In the United States, Fleishman and Cohen (2010) found that medical condition information improved prediction of high expenditures beyond that obtainable using gender and age. Ash and colleagues (2001) also found that risk models based on Diagnostic Cost Groups (DCG) were at least as powerful as prior cost for identifying HCUs. In Ontario, Walraven and colleagues (2010) derived and validated an index to predict early death or unplanned readmission after discharge from hospital. Variables independently associated with this outcome (from which authors derived the mnemonic "LACE") included length of stay ("L"), acuity of the admission ("A"), co-morbidity of the patient (measured with the Charlson co-morbidity index score, "C"); and emergency department use (measured as the number of visits in the six months before admission, "E").

Building upon previous research, we developed a predictive model to identify patients at risk of becoming high-cost healthcare users in Ontario. The methods and results of this predictive modelling are presented here. Potential ways to utilize this information in practice and the next steps are also discussed.

Methodology

The purpose of the model presented in this paper is to predict who will or will not become a high-cost healthcare user in the next year, given various patient-level characteristics in the current year and two previous years. The model predicts HCU status in fiscal year (FY) 2010/11 among patients (users of the healthcare system) from FY 2009/10, using patient characteristics from FY 2007/08 through FY 2009/10. The model was validated by applying it to

predict HCU status in FY 2009/10 using patient characteristics from FY 2006/07 through FY 2008/09 (out-of-sample prediction power). Data for the analysis were obtained from the Ontario Ministry of Health and Long-Term Care (MOHLTC).

The cohort of patients incorporated in the model included all Ontario residents serviced by the Ontario healthcare system during FY 2009/10 in one of the following care types (database in parentheses):

- Physician services – OHIP (Claims History Database)
- Acute in-patient care – AIP (Discharge Abstract Database)
- Day surgery – DS (National Ambulatory Care Reporting System)
- Emergency room – ER (National Ambulatory Care Reporting System)
- Dialysis – (National Ambulatory Care Reporting System)
- Oncology – (National Ambulatory Care Reporting System)
- Outpatient clinic – (National Ambulatory Care Reporting System)
- Rehabilitation – Rehab (National Rehabilitation System)
- In-patient mental health – MH (Ontario Mental Health Reporting System)
- Complex continuing care – CCC (Continuing Care Reporting System)
- Long-term care – LTC (Continuing Care Reporting System)
- Home care – HC (Home Care Database)

Data from various administrative sources were linked using encrypted health insurance numbers. In Ontario, which has a single-payer government health system (OHIP), all patients have unique health insurance numbers that are recorded in all sectors whenever a patient receives a health service.

Patients were excluded if they died during FY 2009/10 (as they could not become high users the next year) or were under five years of age in FY 2009/10 (as the history of disease and health utilization progression is required to build a good predictive model). The validation patient cohort was built similarly based on FY 2008/09 data.

HCUs of healthcare in FY 2010/11 (for the modelling cohort) or in FY 2009/10 (for the validation cohort) were defined as the top 5% cost-incurring healthcare users. In order to identify HCU status, we summed costs across care types for each user. Patient costs for AIP, ER, DS, Rehab, CCC, MH and HC were derived from actual unit cost[1] (actual cost per weighted case) times weighted volume of services (number of weighted cases). Cost for OHIP claims was represented by fees approved. Patient cost for LTC was estimated using average cost per patient per day times patient length of stay. Costs for outpatient oncology, outpatient dialysis and outpatient clinic were not included owing to data quality issues with case mix and cost data in these sectors, and the Ministry's general recommendation not to use these data in funding formulas in Ontario. Users (patients) were sorted in descending order of total expenditures, and the top 5% of users were classified as high-cost healthcare users. A binary variable was created and added to the data to identify patients as either high users or not.

Information was collected on factors (covariates) that may have an influence on the outcome (becoming a high user). These included demographic variables (e.g., age, sex, Rurality Index of Ontario [RIO] Score by the Ontario Medical Association[2]); clinical variables (e.g., ICD-10 based chapters created from ICD-10 and ICD-9 diagnoses[3]), with further separation of certain chronic conditions such as diabetes, congestive heart failure (CHF) and chronic obstructive pulmonary disease (COPD)[4]; socio-economic status (SES) variables (e.g., material and social deprivation indices[5]); and utilization variables for all care types from current year and previous two years, enabling us to account for disease progression (e.g., number of visits, number of hospitalizations).

Continuous variables, if necessary, were transformed to categorical variables (e.g., age into age groups). Continuous variables for healthcare utilization were categorized based on percentiles (zero category was created first, then median was calculated for the remaining positive values, and two remaining categories were created: less than median, and equal or more than median). Where applicable, missing values were imputed via the multiple imputation technique (using SAS PROC MI, SAS Institute). The final data set had missing values for only three variables in our model: Rurality Index of Ontario (1.51% of patients with missing values) and social and material deprivation (3.72% of patients with missing values). The pattern of the missing data allowed us to assume that the values were missing completely at random (MCAR). As a next step, a number of variables were reduced using the variable clustering technique (SAS PROC VARCLUS, SAS Institute).

A logistic regression model predicting the next year's HCU status was built and executed in SAS 9.1.3 on the FY 2009/10 patient cohort. Performance of the model was evaluated using C-statistic for predictive ability of the model. Significance of the parameter estimates (p-values) and odds ratios were evaluated as well.

Validation of the model on the FY 2008/09 cohort was done to evaluate the out-of-sample prediction power. Good predictive models should show strong performance in the new (out-of-sample, scored) data, since in-sample performance could be unduly optimistic if the model over-fitted the data (in this case, out-of-sample performance could be very poor). Moreover, this is the intended application of predictive models: to apply the model to the new data with unknown outcomes in order to predict them. Receiver-operating characteristic (ROC) curve and calibration (goodness-of-fit) curves were constructed based on the validation sample. Out-of-sample model performance was evaluated using sensitivity, specificity, positive and negative predictive values, and accuracy on the validation sample (calculated for scenarios if following up on the top 1%, 5%, 10% or 15% of patients with the highest risk of becoming HCUs). Cut-off probability levels for different potential follow-up cohorts (if following up 1% of highest risk users, 5%, 10% or 15%) were selected, and the outcome variable was set at 1 if the predicted probability equaled or exceeded that cut-off. Multiple cut-offs for follow-up are presented (following up 1%, 5%, 10%, 15%) to enable dialogue on the degree of sensitivity that could be achieved if different resources are utilized (assuming that implementation of follow-up interventions requires additional resources).

Formal ethics review was not required because de-identified Ministry administrative data were used.

Results

Patient characteristics of the model cohort and the validation cohort are presented in Table 1. The population of patients in the FY 2009/10 model cohort was 10,300,856. The number of variables in the initial model was 97. Sixty-four variables were transformed. The number of variables removed because of clustering was 28, leaving 69 variables in the final model. See Table 1 at www.longwoods.com/content/23710.

The model achieved a very strong C-statistic: .865.[6] Odds ratios for all predictor variables used and their 95% confidence intervals are reported in Table 2. Among the patients that the model predicted to be high-cost healthcare users in 2010, 46.2% were not HCUs in 2009. See Table 2 at www.longwoods.com/content/23710.

Odds ratios analysis reveals that age is a strong predictor of becoming a high-cost healthcare user, and there is a clear pattern of substantially increasing risk as age increases. Similarly, as the material and social deprivation indices increase, the risk of becoming a high-cost user increases. Interestingly, social deprivation seems to increase risk more than material deprivation. Males have an increased risk compared to females. There is also a clear pattern of increased risk as the degree of rurality increases (as measured by the Rurality Index of Ontario Score). Current and past (1 year ago and 2 years ago) healthcare utilization across different care types are among the strongest predictors of becoming high-cost healthcare users. Of particular note are long-stay, long-term care utilization, more than one hospitalization in in-patient mental health, chronic continuing care, acute in-patient care, high number of outpatient dialysis and oncology visits, and high number of services in home care. The most influential diagnoses (controlling for all other variables in the model) are mental and behavioural disorders; congestive heart failure, chronic obstructive pulmonary disease and diabetes; diagnoses in pregnancy, childbirth and the puerperium ICD chapter; and congenital malformations, deformations and chromosomal abnormalities.

Based on out-of-sample validation, both the ROC curve (Figure 1) and the calibration (goodness-of-fit) curves (Figure 2) show very good out-of-sample model performance. Table 3 presents sensitivity, specificity, positive and negative predictive values, and accuracy for different cut-off points for the validation (out-of-sample) cohort. If the top 5% patients at risk of becoming HCUs are followed, the achieved sensitivity and specificity is 42.2% and 97%, respectively. These values suggest very reasonable predictive power, indicating that the model picks up 42.2% of all high-cost healthcare users and correctly identifies 97% of those who are not high users. Accuracy of 94.2% is also very reasonable (percentage of true positive and true negative out of all patients).

FIGURE 1. Receiver operating characteristic (ROC) plot of model performance on scored 2008 data

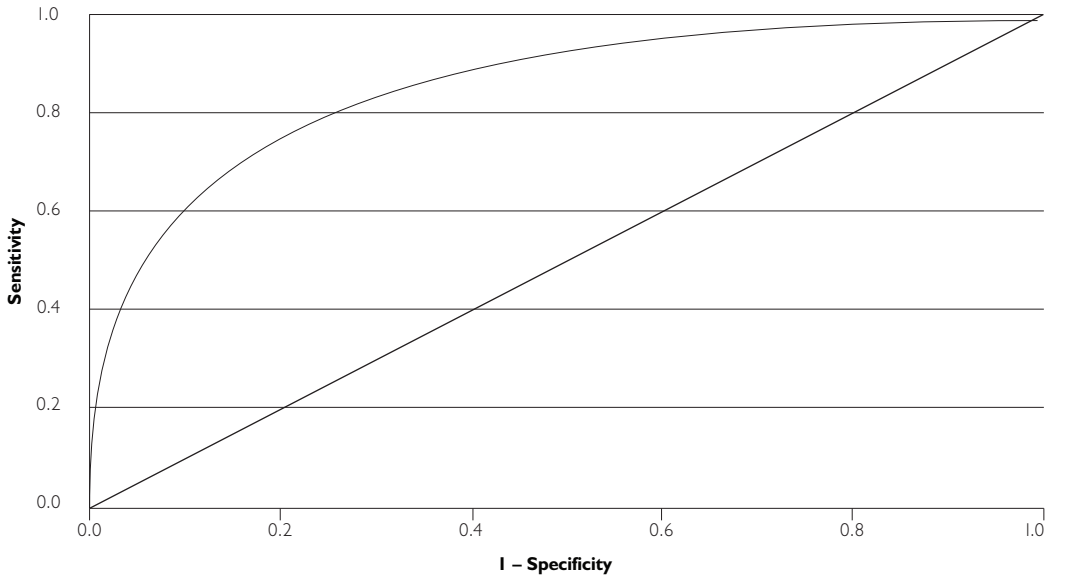

FIGURE 2. Goodness of fit (calibration) curve on scored 2008 data

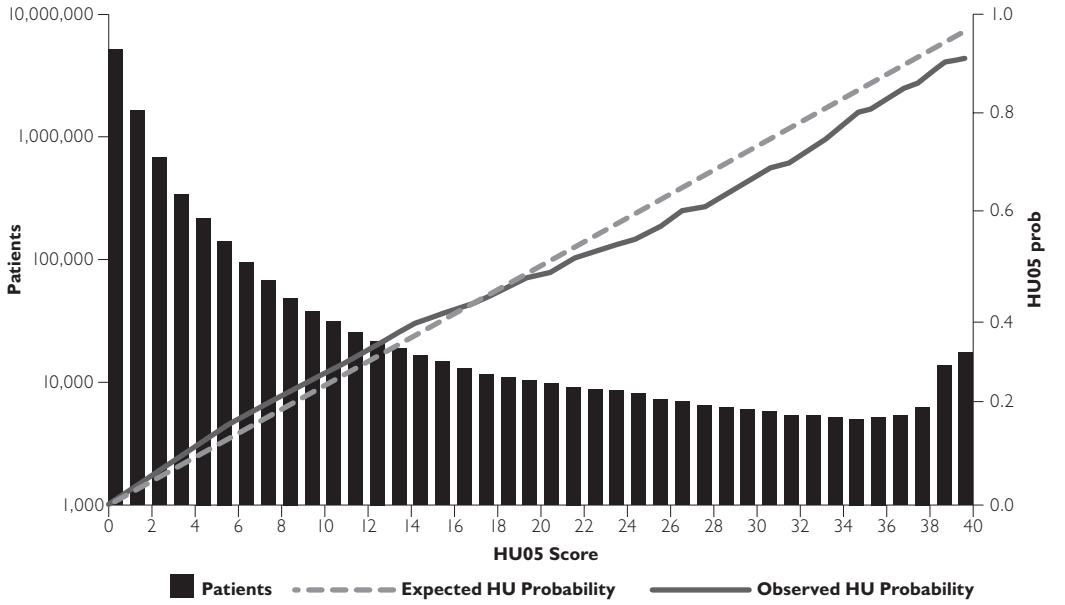

Table 3 also presents sensitivity, specificity and so on for other follow-up cut-off points (such as 1%, 10% and 15%). This analysis is still based on the 5% HCU model (i.e., predicting the risk of becoming a top 5% high-cost healthcare user). However, the cut-off point for follow-up of patients could be selected at any level (following up 1%, or even 10% of patients with the highest risk of becoming top 5% healthcare users), depending on availability of healthcare funding for follow-up, thus increasing or decreasing the sensitivity, or the number of captured future top 5% high-cost healthcare users.

TABLE 3. Predictive (out-of-sample) performance of the model

Metric	Selection of patients based on predicted probabilities – the top:				Formula	Notes
	1%	5%	10%	15%		
Sensitivity	15.8%	42.2%	57.1%	66.4%	TP/(TP + FN)	picks up % of all high users
Specificity	99.8%	97.0%	92.5%	87.7%	TN/(FP + TN)	correctly identifies % of those who are not high users
Positive Predictive Value	79.9%	42.6%	28.8%	22.4%	TP/(TP + FP)	good at confirming high users
Negative Predictive Value	95.7%	96.9%	97.6%	98.0%	TN/(FN + TN)	reassuring that a patient will not become a high user
Accuracy	95.5%	94.2%	90.7%	86.7%	(TP + TN)/(P + N)	% of true positive and true negative out of all patients

Notes: TP – true positive, FN – false negative, TN – true negative, FP – false positive, P – positive, N – negative

Conclusion

This is the first attempt in Ontario to develop and validate a tool for predicting patients at risk of becoming high-cost healthcare users. Presented results suggest that the performance of the model is very good, and the model has been validated for an out-of-sample validation cohort. Special attention could be paid to the factors, listed in the Results section, that are the strongest predictors of becoming HCUs.

The performance of this model (in terms of sensitivity, for example) in comparison to other published predictive models could be of interest. Unfortunately, the authors were not able to identify published studies with reported performance for models predicting high-cost healthcare users. Nevertheless, we scanned the literature to observe the performance of other predictive models with relevant outcomes that could be an indirect proxy for future high cost, such as hospital admissions in the next year. The literature scan showed that our model has similar or better performance. For example, SPARRA (Scottish Patients At Risk of Readmission and Admission) (Government of Scotland 2012), a model that predicts next-year hospital admissions, reports 10.5% sensitivity at the 50% risk threshold (which corresponds to the top 1.6% patients follow-up), and the best performing model, IPAEOPGP (using in-patient, accident and emergency, and outpatient data, and general practitioners' electronic medical records data) (Billings et al. 2013), which also predicted next-year admissions, achieved 9.2% sensitivity for the top 1% patients follow-up. In comparison,

our model demonstrated 15.8% sensitivity for the top 1% patients follow-up. The same SPARRA model reports approximately 27% sensitivity at the 30% risk threshold (corresponding to the top 5.5% patients follow-up), and the IPAEOPGP model reports 28.5% sensitivity for the top 5% patients follow-up. In comparison, our model's sensitivity is 42.2% for the top 5% patients follow-up. While these numbers are not precisely comparable – because the models predict different outcomes, and these outcomes have different event rates (prevalence) – they provide overall reference points for our predictive model's performance. It should be noted that our model reports performance in the validation cohort (out-of-sample model performance). Table 4 summarizes the information on the performance of the models discussed, based on sensitivity and positive predictive value metrics.

TABLE 4. Comparison of model performance

Selection of patients based on predicted probabilities	Performance metric	Model***		
		SPARRA	IPAEOPGP	Current
Top 1%*	Sensitivity	10.5%	9.2%	15.8%
	Positive Predictive Value	59.8%	47.5%	79.9%
Top 5%**	Sensitivity	27.0%	28.5%	42.2%
	Positive Predictive Value	44.1%	29.4%	42.6%

* For SPARRA model it is top 1.6%

** For SPARRA model it is top 5.5%

*** Models compared: SPARRA: SPARRA model (Scottish Patients At Risk of Readmission and Admission), predicting next-year hospital admissions (Government of Scotland 2012); IPAEOPGP: IPAEOPGP model (using in-patient, accident and emergency, and outpatient data, and general practitioners' electronic medical records data, predicting next-year hospital admissions) (Billings et al. 2013)

Current: Our model, predicting high-cost healthcare users

Please note that these numbers are not precisely comparable because the models predict different outcomes and have different event rates (prevalence); however, the figures provide overall reference points for our predictive model's performance.

This paper has presented a model that predicts the top 5% high-cost healthcare users. Models also predicting the top 1% and top 10% HCUs were also explored using reported methodology, and these models showed very strong performance. The final decision as to the most useful cut-off point will depend on the specifics of policy decision-making with regard to practical implementation, and of course on the availability of scarce healthcare dollars for the follow-up.

Limitations of the current model include a very large number of predictor variables and the heavy data requirements to run the model. The number of variables could potentially be further reduced in the future. Missing values do not present a significant obstacle, as only three variables have missing values for a very small proportion of patients. Another limitation that we encountered was an inability to access available patient classification systems (such as Adjusted Clinical Groups [ACG] or Diagnostic Cost Groups [DCG]). Usage of such systems

could further improve model performance. In order to overcome this obstacle, we used a proxy in our model: all patient ICD-10 and ICD-9 diagnoses were grouped into ICD-10 chapters, with further separation of certain chronic diseases, such as CHF, COPD and diabetes. Odds ratios showed that these groups were very strong predictors of future HCU status.

Some elements of our model are being used by the Ontario Ministry of Health and Long-Term Care to support Health Links, a new model of care at the clinical level where all providers in a community – including primary, hospital and community care – are involved in coordinating plans at the patient level. The Ministry's Health Analytics Branch develops community profiles of populations and high users where the Health Links model is being established.

Further practical implementation of the model could occur in a number of ways. One potential approach is to provide the health card numbers of high-risk patients to primary care providers so that they can implement appropriate prevention strategies, potentially mitigating or avoiding HCU status in the future. For example, the Government of Scotland (2010) has described a model in which lists of patients at greatest risk of emergency admission to hospital over the next year are distributed to healthcare providers to enable "delivery of Proactive, Planned and Co-ordinated care for people with complex or frequently changing care needs."

In Ontario, there are several challenges to this approach, including concerns over privacy in regard to sharing data with providers. The Ministry's approach with Health Links[7] has been to provide aggregate information about HCUs so that providers can identify the patient populations that have historically consumed the most resources. This model could be used to identify patient populations at high risk of becoming high users, so that Health Links could develop interventions that address specific needs of those patient groups. An example might be CHF-centred clinics for congestive heart failure patients (Wijeysundera et al. 2012). Another potential approach would require converting this data-intensive model into a simpler one, with fewer variables, and creating a paper- or desktop computer-based tool that can be used by providers themselves in their offices to score patients and identify those at high risk of becoming high users.[8] In any scenario, physicians would be informed of at-risk patients to provide timely interventions to mitigate or reduce the number of HCUs, thus improving patient outcomes and saving finite resources. Ontario's Action Plan for Health Care (MOHLTC n.d.) calls for "better patient care through better value from our health care dollars," and the current study could become one of the tools facilitating implementation of the directions identified in the Plan.

Correspondence may be directed to: Yuriy Chechulin, Senior Methodologist, Methods and Modelling, Health Analytics Branch, Ministry of Health and Long-Term Care, 1075 Bay Street, 13th floor, Toronto, ON, M5S2B1; tel.: 416-327-9309; e-mail: yuriy.chechulin@ontario.ca.

Notes

1. Actual costs in Ontario are established for each hospital using Ontario Cost Distribution Methodology (OCDM), Ontario Ministry of Health and Long-Term Care.
2. Rurality Index of Ontario methodology is described in the following document: http://www.health.gov.on.ca/english/providers/program/uap/docs/up_rio_methodology.pdf.
3. Diagnosis codes for patients were collected from all care types (all healthcare encounters) for the whole year. All diagnoses listed on a patient record were used, not just the principal diagnosis.
4. Owing to lack of access to population classification systems (such as ACG, etc.; see detailed discussion in the Conclusion section on model limitations), we used ICD-10 and ICD-9 diagnoses grouped into ICD-10 chapters with further separation of certain chronic conditions (such as CHF, COPD, diabetes). Dummy variables were entered in the model specifying whether a patient has or does not have a disease (ICD-10 or ICD-9 code) in the corresponding ICD-10 chapter (ICD-9 chapters were mapped onto ICD-10 chapters).
5. Material deprivation primarily portrays variations associated with education, employment and income. Social deprivation indicates the state of being separated, divorced or widowed, living alone, or being a member of a single-parent family (Pampalon et al. 2009).
6. C-statistic is the probability that predicting the outcome is better than chance. It is used to compare the goodness of fit of logistic regression models. Values for this measure range from 0.5 to 1. A value of 0.5 indicates that the model is no better than chance at making a prediction of membership in a group; a value of 1 indicates that the model perfectly identifies those within a group and those not. Models are typically considered reasonable when the C-statistic is higher than 0.7 and strong when C exceeds 0.8 (Hosmer and Lemeshow 2000).
7. Ontario's Health Links initiative aims to facilitate coordination of care at a local level for high-needs patients. One of the goals of the initiative is to provide better care for the 1% to 5% of people who, research has indicated, are high users of healthcare. It also aims to reduce costs, particularly expensive hospital visits, based on the assumption that many of these patients' hospital emergency ward visits, admissions and readmissions can be prevented with better coordinated care (Silversides and Laupacis 2013).
8. The authors are currently working on developing such a model.

References

Ash, A.S., Y. Zhao, R.P. Ellis and M.S. Kramer. 2001. "Finding Future High-Cost Cases: Comparing Prior Cost versus Diagnosis-Based Methods." *Health Services Research* 36(6 Part II): 194–206.

Berk, M.L. and A.C. Monheit. 2001. "The Concentration of Health Care Expenditures, Revisited." *Health Affairs* 20(2): 9–18.

Billings, J., J. Dixon, T. Mijanovich and D. Wennberg. 2006. "Case Finding for Patients at Risk of Readmission to Hospital: Development of Algorithm to Identify High-Risk Patients." *British Medical Journal* 333(7563): 327. doi: 10.1136/bmj.38870.657917.AE.

Billings, J., T. Georghiou, I. Blunt and M. Bardsley. 2013. "Choosing a Model to Predict Hospital Admission: An Observational Study of New Variants of Predictive Models for Case Finding." *BMJ* Open 3: e003352. doi:10.1136/bmjopen-2013-003352.

Calver, J., K.J. Brameld, D.B. Preen, S.J. Alexia, D.P. Boldy and K.A. McCaul. 2006. "High-Cost Users of Hospital Beds in Western Australia: A Population-Based Record Linkage Study." *Medical Journal of Australia* 184: 393–97.

Center for Healthcare Research and Transformation. 2010. "Health Care Cost Drivers: Chronic Disease, Comorbidity, and Health Risk Factors in the U.S. and Michigan." Retrieved January 29, 2014. <http://www.chrt.org/assets/price-of-care/CHRT-Issue-Brief-August-2010.pdf>.

Fleishman, H.A. and J.W. Cohen. 2010. "Using Information on Clinical Conditions to Predict High-Cost Patients." *Health Services Research* 45(2): 532–52.

Gawande, A. 2011 (January). "The Hot Spotters: Can We Lower Medical Costs by Giving the Neediest Patients Better Care?" *The New Yorker*. January 2011. Retrieved January 29, 2014. <http://www.newyorker.com/reporting/2011/01/24/110124fa_fact_gawande>.

Government of Scotland. 2010. "SPARRA Made Easy." Retrieved January 29, 2014. <http://www.scotland.gov.uk/Publications/2010/04/13104303/1>.

Government of Scotland. 2012. "Scottish Patients at Risk of Readmission and Admission (SPARRA): A Report on the Development of SPARRA Version 3 (Developing Risk Prediction to Support Preventative and Anticipatory Care in Scotland)." Retrieved January 29, 2014. <http://www.isdscotland.org/Health-Topics/Health-and-Social-Community-Care/SPARRA/2012-02-09-SPARRA-Version-3.pdf>.

Hosmer, D.W. and S. Lemeshow. 2000. *Applied Logistic Regression* (2nd ed.). New York: John Wiley & Sons.

Kozyrskyj, A., L. Lix, M. Dahl and R. Soodeen. 2005 (March). *High-Cost Users of Pharmaceuticals: Who Are They?* Winnipeg: Manitoba Centre for Health Policy. Retrieved January 29, 2014. <http://mchp-appserv.cpe.umanitoba.ca/reference/high-cost.pdf>.

Ministry of Health and Long-Term Care (MOHLTC). n.d. "Ontario's Action Plan for Health Care." Retrieved January 29, 2014. <http://www.health.gov.on.ca/en/ms/ecfa/healthy_change/docs/rep_healthychange.pdf>.

Moturu, S.T., W.G. Johnson and H. Liu. 2010. "Predictive Risk Modelling for Forecasting High-Cost Patients: A Real-World Application Using Medicaid Data." *International Journal of Biomedical Engineering and Technology* 3(1/2): 114–32.

Pampalon, R., D. Hamel, P. Gamache and G. Raymond. 2009. "A Deprivation Index for Health Planning in Canada." *Chronic Diseases in Canada* 29(4): 178–91.

Rais, S., A. Nazerian, S. Ardal, Y. Chechulin, N. Bains and K. Malikov. 2013. "High Cost Users of Ontario's Healthcare Services." *Healthcare Policy* 9(1): 44–51.

Reid, R., R. Evans, M. Barer, S. Sheps, K. Kerluke, K. McGrail, C. Hertzman and N. Pagliccia. 2003. "Conspicuous Consumption: Characterizing High Users of Physician Services in One Canadian Province." *Journal of Health Services Research & Policy* 8(4): 215–24.

Silversides, A. and A. Laupacis. 2013 (February 28). "Health Links: Ontario's Bid to Provide More Efficient and Effective Care for Its Sickest Citizens." Retrieved January 29, 2014. <http://healthydebate.ca/2013/02/topic/innovation/the-ontario-health-links-initiative-what-is-it>.

The Commonwealth Fund. 2012 (April). "The Performance Improvement Imperative: Utilizing a Coordinated, Community-Based Approach to Enhance Care and Lower Costs for Chronically Ill Patients." Retrieved January 29, 2014. <http://www.commonwealthfund.org/Publications/Fund-Reports/2012/Apr/Performance-Improvement-Imperative.aspx>.

Walraven, C., D.A. Irfan, C. Bell, E. Etchells, I.G. Stiell, K. Zarnke et al. 2010. "Derivation and Validation of an Index to Predict Early Death or Unplanned Readmission After Discharge from Hospital to the Community." *Canadian Medical Association Journal* 182(6). doi: 10.1503/cmaj.091117.

Wijeysundera, H.C., G. Trubiani, L. Abrahamyan, N. Mitsakakis, W. Witteman, M. Paulden et al. 2012. "Specialized Multi-Disciplinary Heart Failure Clinics in Ontario, Canada: An Environmental Scan." *BMC Health Services Research* 12: 236. doi: 10.1186/1472-6963-12-236.