EARLY IDENTIFICATION OF PEOPLE AT-RISK OF HOSPITALIZATION

Hospital Admission Risk Prediction (HARP) – a new tool for supporting providers and patients
## Table of Contents

**Introduction** ......................................................... 4  
  Early intervention ............................................... 4  
  Predictive modelling ............................................. 4  
  Supporting providers ............................................. 5

**Methods** ............................................................... 6  
  Stage 1 – Variable selection ........................................ 6  
  Stage 2 – Algorithm derivation and validation ................. 7

**Results** ................................................................. 8  
  Variable selection results ........................................ 8  
  Algorithm validation results ..................................... 8  
  Figure 1: Simple algorithm ....................................... 8  
  Figure 2: Complex algorithm ..................................... 8  
  Figure 3: Calibration curve of the simple algorithm for admission within 30 days based on validation data .......... 9  
  Figure 4: HARP algorithm statistics .................. 10

**Discussion** ............................................................. 11  
  Selecting a model ................................................... 11

**Looking ahead** ....................................................... 12  
  HARP in action ..................................................... 12  
  Continuing analysis ............................................... 12  
  Cost effective care ............................................... 12

**Acknowledgments** .................................................. 13

**References** ........................................................... 14
Executive Summary

The burden of disease has a major impact on the health and lives of Ontarians. It also exerts a great cost on the health system. An estimated five percent of Ontario’s population accounts for nearly 85 percent of total provincial spending on combined hospital and home care costs; one percent of Ontarians account for nearly half of that spending.

Early intervention can relieve some of the burden of illness from patients and free up strained system resources, while reducing in-patient costs. The key to effective early intervention is identifying the patients most likely to benefit, before they require multiple hospitalizations and become high users of the health system. Predictive risk models are a relatively new and effective means to support care providers in identifying these patients.

Drawing on lessons from other existing models and evidence, Health Quality Ontario (HQO), the Canadian Institute of Health Information (CIHI), and other experts partnered to design a model that would identify an individual patient’s near and longer-term risk of future hospitalization.

The Hospital Admission Risk Prediction (HARP) tool accounts for a number of patient-level variables that are predictive of future hospitalization. HARP generates an individual patient risk score of hospital admission within two timeframes: 30 days, and 15 months. Its use can improve the quality of targeted care for at-risk patients, and generate a savings for our health care system overall when used for early, cost-effective intervention.

The variables most predictive of future hospitalization, and included in the HARP tool are: the patient’s age, the number of admissions and emergency department visits in the past six months, location where the patient was previously discharged to, the intensity of a previous admission, the presence of the 18 top conditions, whether a previous admission was through the emergency department, the Charlson co-morbidity index, select interventions during a hospital encounter, and previous length of stay. The simplified model omits much of the hospital data, and accounts for the top six conditions.

This report describes the development of the HARP tool, including the companion simple model, and the findings of validation analyses. Aside from producing a functional risk model, the work showed that community characteristics are not predictive of future hospitalization. The validation analysis identified that the simple model is a strong substitute for the complex model that accounts for a wider set of factors. Future work, including the incorporation of primary care factors, and the promotion and use of HARP among practitioners is also discussed.
Introduction

The burden of disease greatly affects Ontarians’ lives and strains available health care resources. Early intervention can benefit ill patients and generate system savings when care is appropriately targeted to those most at-risk. That targeted focus is a key challenge to improving care and supporting health care providers in caring for their patients. Predictive modelling has been a useful tool to define high-risk patients and is the focus of the work discussed in this paper.

Early intervention

In 2009-10, approximately 34,100 people were admitted to hospitals in Ontario for complications from chronic disease that could potentially have been prevented with improved primary care, or early intervention. These diseases are a great burden on the health system and on people’s lives. Chronic diseases are the leading cause of death in Ontario and diminish our quality of life, our economy, and the well-being of our communities.

Not only does the burden of disease have a major impact on the health and lives of Ontarians, it also exerts a huge cost. An estimated five percent of Ontario’s population accounts for nearly 85 percent of total provincial spending on combined hospital and home care costs; one percent of Ontarians account for nearly half of that spending. Just 39,000 very high users account for over 30% of hospital and home care costs.

There is a growing body of evidence that early interventions targeted to those most at-risk of illness hold promise in avoiding unplanned admissions, reducing emergency department congestion, and reducing in-patient costs. It has been forecasted that a five percent cost reduction for the top five percent of patients who are high users of the health care system would save Ontario $760 million in health care costs. A 10 percent reduction could save our system $1.5 billion. For the top one percent of high-cost patients, the same reduction has the potential to generate $785 million in savings. A 15 percent reduction could amount to nearly $1.2 billion in savings.

Other jurisdictions have had success in designing interventions that reduce avoidable hospitalizations and readmissions. One promising example is an American model of enhanced primary care and care coordination that has decreased hospitalizations by 43 percent and emergency department visits by 70 percent. A chronic care coordination model operated by Kaiser Permanente has also shown success, decreasing re-hospitalizations from 14 percent to 2.4 percent. In Ontario, the “virtual ward” model is in use in Toronto to provide intensive management in the community for patients at high risk for readmission to acute care.

Predictive modelling

A key component of an early intervention strategy is the identification of patients most “at-risk” of hospitalization. When interventions are targeted to those most likely to become high users, the likelihood of success and cost savings can increase dramatically. Health providers and planners should be aware of the tendency for patient admissions to reach a peak, and then naturally decrease over time, even without intervention; a phenomenon known as “regression to the mean”. Identifying patients prior to that peak point in their illness is important for avoiding health deterioration and for structuring cost-effective programs that can truly reduce avoidable admissions.

Predictive risk modeling using regression analysis can help providers identify at-risk, destabilizing patients and can give providers information to trigger earlier care intervention. Predictive models have been shown to be more effective than relying on clinical judgment or standard checklists alone, which tend to identify patients currently at-risk rather than
those at future risk. Predictive ability varies between models and some perform poorly, which speaks to the importance of a careful model construction and rigorous testing. Predictive models are currently used in a number of countries, including the United States and the United Kingdom. In Canada, the Ontario developed LACE (Length of stay, Acuity of admission, patient Comorbidity, and number of visits to the Emergency room) model is being used to predict 30 day readmissions to hospital.

HQO partnered with CIHI and other researchers to develop a model that provided a longer-term assessment of risk of hospitalization than any existing model, in order to support early care intervention in a community setting. The Hospital Admission Risk Prediction (HARP) tool accounts for a number of patient-level variables that are predictive of future hospitalization. HARP generates an individual patient risk score of hospital admission within two timeframes: 30 days, and 15 months.

Supporting providers
Health care providers can integrate the model into existing medical record systems, drawing on at-hand patient-level information. By leveraging this information for earlier care, health care providers can both help their patients prevent disease progression and avoid hospital visits, while avoiding unnecessary health system costs. Without predictive risk models, resource intensive interventions may be less likely to focus on those patients most at-risk for future hospitalizations, and be less cost effective. Thus, through predictive modeling, it is possible to improve the quality of targeted care for at-risk patients, and generate a savings for our health care system overall.

The HARP tool is available for use in two forms: a “simple” model, and a “complex” model. The simple model accounts for five factors, while the complex model accounts for 10 factors. The complex model performs slightly better, but the simple model is a strong substitute for care settings without access to hospital data sources. Both models are able to produce individual risk scores for hospitalization within 30 days or 15 months. The statistical code needed to stratify patient populations according to risk score is now available to health providers and planners, without charge.
Methods

Model development was based on multiple regression analyses to estimate the relationship between patient characteristics and risk for future hospitalization. A two-stage approach was taken for tool development: 1) the identification of variables predictive of hospitalization, and 2) derivation and validation of predictive algorithms.

Stage 1 – Variable selection
The Stage 1 analysis involved selecting variables of interest, defining a test population, running statistical analysis to define an impact of the variables to admission, and selecting two sets of variables – a smaller set for a simple model and a greater set for a complex model.

A wide range of candidate variables were identified for inclusion in the analysis. These included known influencing factors, variables identified in available data sources, and variables used in other known risk models.

The study population used in developing HARP consisted of individuals with an initial hospital stay. Identifying patients at risk in the community, prior to hospital care and using information from primary care encounters and patient medications were of great interest to the group involved. Such information would also help address the “regression to the mean” effect by providing upstream data that can enable earlier care intervention. However, due to data availability limitations, incorporating them into the analysis had to be deferred to a future phase of the project. The variables included in the analysis fell into three broad categories: patient demographic and community characteristics, patient disease and condition, and patient encounters with the hospital system:

- Demographic variables included the patient level variables (age and sex) and community characteristics (rural residence, neighborhood income quintiles, community admission rates, the four dimensions of the Canadian Marginalization Index: residential instability, material deprivation, dependency, and ethnic concentration).

- A number of diseases and conditions, and resource intensity levels were included as candidate variables, using CIHI’s 2012 Case Mix Groups.

- Variables from hospital data included a range of interventions, length of stay, and other episodic service variables.

Data sources for these variables included the Discharge Abstract Database, and the National Ambulatory Care Reporting System.

A database was created of adult patients discharged from hospital following a medical encounter; mental health, obstetric, pediatric, and surgical admissions were not included in the study population as the expert panel determined admissions would in these cases either be planned, elective, or influenced by variables significantly different than the target population. Manitoba data was also included to increase the sample and to later determine the model’s applicability to another Canadian jurisdiction. A total of 385,065 initial, index episodes were identified from discharges in Ontario and Manitoba in 2009-10. Subsequent medical admissions were identified in data from 2009-10 to 2011-2012. History of emergency department use and acute care hospitalization (six months prior to admission) was drawn from 2008-09 and 2009-10 data. Patient co-morbidity was measured using the Charlson index. A multivariate regression analysis was

---

19 The study population used in developing HARP consisted of individuals with an initial hospital stay. Identifying patients at risk in the community, prior to hospital care and using information from primary care encounters and patient medications were of great interest to the group involved. Such information would also help address the “regression to the mean” effect by providing upstream data that can enable earlier care intervention. However, due to data availability limitations, incorporating them into the analysis had to be deferred to a future phase of the project. The variables included in the analysis fell into three broad categories: patient demographic and community characteristics, patient disease and condition, and patient encounters with the hospital system:

- Demographic variables included the patient level variables (age and sex) and community characteristics (rural residence, neighborhood income quintiles, community admission rates, the four dimensions of the Canadian Marginalization Index: residential instability, material deprivation, dependency, and ethnic concentration).

- A number of diseases and conditions, and resource intensity levels were included as candidate variables, using CIHI’s 2012 Case Mix Groups.

- Variables from hospital data included a range of interventions, length of stay, and other episodic service variables.

Data sources for these variables included the Discharge Abstract Database, and the National Ambulatory Care Reporting System.

A database was created of adult patients discharged from hospital following a medical encounter; mental health, obstetric, pediatric, and surgical admissions were not included in the study population as the expert panel determined admissions would in these cases either be planned, elective, or influenced by variables significantly different than the target population. Manitoba data was also included to increase the sample and to later determine the model’s applicability to another Canadian jurisdiction. A total of 385,065 initial, index episodes were identified from discharges in Ontario and Manitoba in 2009-10. Subsequent medical admissions were identified in data from 2009-10 to 2011-2012. History of emergency department use and acute care hospitalization (six months prior to admission) was drawn from 2008-09 and 2009-10 data. Patient co-morbidity was measured using the Charlson index.

---

a Global exclusions included invalid fields (for health care number, age, sex, admission/discharge date/time), death at discharge/leave against advice, mental illness as primary diagnosis, newborn episodes, and chemotherapy/radiotherapy.

b Not available for Manitoba.
then run to determine the relationship between identified variables and influence on future hospital admissions. Three separate analyses were completed with the dependent variable being one, six, and 12 months, respectively. A bivariate analysis was also undertaken to assess the impact of any one individual variable on future admission.

Results were analyzed, expressed as odds-ratios with 95 percent upper and lower limits. C statistics, R squares and goodness of fit statistics were also evaluated for each of the time periods (refer to the Technical Appendix for further detail on statistical tests used in this study). Variables that had no significant impact on future admissions were eliminated, after confirming in the bivariate results. “Simple” and “complex” models were derived to determine the additional benefits available relative to the cost of gathering additional clinical data. This approach also had the benefit of assessing the trade-off between having a simple, easily understood model and the benefit of greater predictive ability.

**Stage 2 – Algorithm derivation and validation**

The Stage 2 analysis included defining the population, deriving the algorithms, weighing the selected variables and validating the algorithm.

For each model, a one month and 15 month algorithm was developed for future risk of admission. A split sample design was taken for algorithm derivation and validation, using an episode of care for the unit of analysis. The study population (382,948 acute medical episodes in Ontario and Manitoba in 2009-10) was split in two parts using random assignment: the derivation dataset (191,321 episodes) and the validation dataset (191,627 episodes). A multivariate regression analysis was again run, with the refreshed list of variables. If any variable was found to be insignificant in the new analysis, the variable was eliminated and the regression results were recalculated.

To assign weighting to variables included in the algorithms, each variable’s correlation coefficient was divided by the smallest coefficient in the model, and then rounded to the nearest whole number – a methodology described by Sullivan and colleagues and employed in the development of the LACE readmission index. C statistics, pseudo R-square, and Hosmer-Lemeshow tests were calculated for each model.

Each algorithm was then run in the validation sample and the expected readmission rate was compared to the observed rate, according to the risk score. The following statistics were also calculated: model sensitivity and specificity, positive and negative predictive values, and kappa coefficients. Following the analysis, the derivation and validation exercises were replicated by a second CIHI analyst.
Results

Variable selection results
The initial logistic regression analysis found a range of clinical and service-delivery factors were predictive of future admissions. The strongest predictors were the location to which the patient was discharged, the patient’s history of hospital and emergency department use, and patient age. Several conditions were predictive of future hospitalization: chronic obstructive pulmonary disease (COPD), heart failure, inflammatory bowel disease (IBD), gastrointestinal (GI) obstruction, and cirrhosis were among the strongest predictors. The Charlson comorbidity index, which predicts mortality for patients with co-morbid conditions,26 was predictive, although the relationship between hospitalization and increasing number of co-morbidities was not strong.

Community characteristics were not predictive, as compared to service variables. Community admission rate was only weakly related to individual risk of admission. Income quintiles and the four indices of the Canadian Marginalization Index (residential instability, material deprivation, dependency, and ethnic concentration) were not predictive.

Algorithm validation results
Variables that were not significantly predictive of future admissions were dropped from the Stage 2 analysis, and the simple and complex models were developed to account for the impact of the remaining variables on future hospitalization.

As described in the Methods section, a “simple” and “complex” model were developed for comparison. The “simple” algorithm captured five variables: the patient’s age, location where the patient was discharged to, the number of admissions and emergency department visits in the past six months, and presence of the six top conditions, ranked by prevalence and predictive strength: COPD, heart failure, IBD, GI obstruction, and cirrhosis. See Figure 1 for all values included in the “simple” model.

Figure 1: Simple algorithm

<table>
<thead>
<tr>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age: 65-84, 85+ vs. other</td>
</tr>
<tr>
<td>Place patient was discharged to: acute (30 day model only), home care vs. other</td>
</tr>
<tr>
<td>Number of acute admissions, six months prior: 1/2/3+ vs. 0</td>
</tr>
<tr>
<td>Number of emergency department visits, six months prior: 1/2/3/4+ vs. 0</td>
</tr>
<tr>
<td>Top Case Mix Groups (CMGs) based on both prevalence and predictive strength: COPD, heart failure, IBD, GI obstruction, cirrhosis</td>
</tr>
</tbody>
</table>

The “complex” algorithm included 10 variables: the five from the simple model, the Resource Intensity Level (a measure of the intensity of resource use),27 whether there was an admission through the emergency department, the Charlson index, select interventions during hospital encounter, and length of stay. The number of conditions in the complex model expanded to 18 conditions, with the highest odds ratios. See Figure 2 for all values included in the “complex” model.

Figure 2: Complex algorithm

<table>
<thead>
<tr>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factors included in simple algorithm, plus:</td>
</tr>
<tr>
<td>Resource intensity level: 2/3/4 vs. 1</td>
</tr>
<tr>
<td>Length of stay (15 month model only): 3-7/8-14/15-30/30+ vs. 0-2 days</td>
</tr>
<tr>
<td>Admission via emergency department: yes vs. no</td>
</tr>
<tr>
<td>Charlson index for co-morbidities: 1/2/3+ vs. 0</td>
</tr>
<tr>
<td>Paracentesis (30 day model only): yes vs. no</td>
</tr>
<tr>
<td>18 CMGs with high OR: refer to technical appendix for full listing</td>
</tr>
</tbody>
</table>
The algorithms were found to be highly concordant in the validation sample. See Figure 3 for agreement in the simple model at 30 days risk of readmission, and the Technical Appendix for graphs of all algorithms. Similar results were found when testing the algorithm against Ontario-only data and in Manitoba-only data, although confidence intervals were much wider in the latter due to a smaller dataset. This concordance suggests that HARP could be of pan-Canadian relevance.

**Figure 3: Calibration curve of the simple algorithm for admission within 30 days based on validation data**

![Calibration Curve](image)

**Notes**

The results are based on Ontario and Manitoba.

**Sources**


The following tables in Figure 4 show results for each of the four models, at the 50th, 75th, and 90th percentiles of risk. As an example of the positive predictive value (PV+), the simple model correctly identified those that would be admitted to hospital in the following 30 days 29% of the time, in the highest risk bracket. For reference, the event rate shows that 12.9% of the study population would be hospitalized in the same time period indicating that the simple model was more than twice as good as random chance in identifying a readmission within 30 days. In all cases, sensitivity decreases and specificity and positive predictive value increase at higher percentiles of the predicted risk score. The sensitivity decreases substantially in all models as there are far fewer events to predict at higher risk scores and many people at low risk are readmitted for reasons not identified by the predicted risk score. Overall the model performance is only slightly better for the complex algorithm over the simple algorithm.
### RESULTS

#### Simple algorithm for 30-day admission

<table>
<thead>
<tr>
<th></th>
<th>50th percentile</th>
<th>75th percentile</th>
<th>90th percentile</th>
<th>Event rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of general population</td>
<td>1.5%</td>
<td>0.8%</td>
<td>0.3%</td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>75.28%</td>
<td>48.82%</td>
<td>24.12%</td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>45.71%</td>
<td>74.34%</td>
<td>91.39%</td>
<td></td>
</tr>
<tr>
<td>PV+</td>
<td>17.08%</td>
<td>22.04%</td>
<td>29.38%</td>
<td>12.94%</td>
</tr>
<tr>
<td>PV-</td>
<td>92.56%</td>
<td>90.72%</td>
<td>89.02%</td>
<td>87.06%</td>
</tr>
<tr>
<td>Kappa</td>
<td>0.09</td>
<td>0.15</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>C statistic:</td>
<td>0.661</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Complex algorithm for 30-day admission

<table>
<thead>
<tr>
<th></th>
<th>50th percentile</th>
<th>75th percentile</th>
<th>90th percentile</th>
<th>Event rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of general population</td>
<td>1.5%</td>
<td>0.8%</td>
<td>0.3%</td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>74.85%</td>
<td>50.65%</td>
<td>24.26%</td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>49.82%</td>
<td>74.63%</td>
<td>91.45%</td>
<td></td>
</tr>
<tr>
<td>PV+</td>
<td>18.14%</td>
<td>22.88%</td>
<td>29.65%</td>
<td>12.94%</td>
</tr>
<tr>
<td>PV-</td>
<td>93.02%</td>
<td>91.05%</td>
<td>89.04%</td>
<td>87.06%</td>
</tr>
<tr>
<td>Kappa</td>
<td>0.11</td>
<td>0.17</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>C statistic:</td>
<td>0.678</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Simple algorithm for 15-month admission

<table>
<thead>
<tr>
<th></th>
<th>50th percentile</th>
<th>75th percentile</th>
<th>90th percentile</th>
<th>Event rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of general population</td>
<td>1.5%</td>
<td>0.8%</td>
<td>0.3%</td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>67.55%</td>
<td>49.57%</td>
<td>20.19%</td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>58.46%</td>
<td>75.61%</td>
<td>93.69%</td>
<td></td>
</tr>
<tr>
<td>PV+</td>
<td>51.48%</td>
<td>57.01%</td>
<td>67.62%</td>
<td>39.49%</td>
</tr>
<tr>
<td>PV-</td>
<td>73.41%</td>
<td>69.68%</td>
<td>64.27%</td>
<td>60.51%</td>
</tr>
<tr>
<td>Kappa</td>
<td>0.25</td>
<td>0.26</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>C statistic:</td>
<td>0.687</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Complex algorithm for 15-month admission

<table>
<thead>
<tr>
<th></th>
<th>50th percentile</th>
<th>75th percentile</th>
<th>90th percentile</th>
<th>Event rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of general population</td>
<td>1.5%</td>
<td>0.8%</td>
<td>0.3%</td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>69.88%</td>
<td>42.52%</td>
<td>20.15%</td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>58.50%</td>
<td>81.93%</td>
<td>93.74%</td>
<td></td>
</tr>
<tr>
<td>PV+</td>
<td>52.41%</td>
<td>60.55%</td>
<td>67.76%</td>
<td>39.49%</td>
</tr>
<tr>
<td>PV-</td>
<td>74.88%</td>
<td>68.59%</td>
<td>64.27%</td>
<td>60.51%</td>
</tr>
<tr>
<td>Kappa</td>
<td>0.27</td>
<td>0.26</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>C statistic:</td>
<td>0.702</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Sources:**

Discharge Abstract Database 2008-2009 to 2011-2012, and National Ambulatory Care Reporting System 2008-2009 to 2009-2010, Canadian Institute for Health Information
Discussion

The HARP predictive models can assist health providers and planners identify patients at-risk of future hospitalization and intervene in a timely manner.

The rigorous approach taken to develop multiple algorithms facilitated the comparison of relative predictive ability and assessment of the trade-offs between simplicity over complexity, in model comprehension and predictive power. Overall, the algorithms show moderate discriminative ability, according to standards used by Kansagara, et al.\textsuperscript{28}

Selecting a model

The expert committee agreed that high sensitivity (correctly predicting hospitalization) is arguably the most desirable when using a model to identify patients most likely to be at a heightened risk. Using that metric, the simple and complex algorithms are quite closely aligned, with respect to performance in predicting future hospitalization. When considering the goodness-of-fit statistics, the complex algorithms slightly outperform.

It is recommended that, where possible, the complex algorithm be used to predict the risk of a patient requiring hospitalization. The ability to use these algorithms may be dependent on the place-of-care’s data environment; the simple model may be simpler to apply outside of a hospital setting as it does not require the data on such variables as hospital intervention, Charlson index of comorbidities, or resource intensity levels.

Overall, the simple model is a strong substitute for the complex model.

Using HARP: A practical example of early intervention

• Through HARP, a primary care provider is alerted that Ms. Jones has been flagged as high-risk for an admission within the next year. She has not visited the clinic in the past several months, so an appointment is arranged. During the visit, she and the care team agree that she requires more intensive support and she is referred to a community program resourced to provide care at her home on a frequent basis.
Looking ahead

HARP in action
HQO will lead the dissemination of HARP by encouraging and supporting its use in communities of practice keen to identify patients at-risk of hospitalization. Both acute-based and community programs will be supported, as will new innovative care models. The model is also expected to assist Ontario’s new Health Links communities in coordinating care for high needs patients. CIHI’s Voluntary Reporting System has been identified as an information platform for seamlessly generating and displaying patient risk information for participating primary care practices. Resource guides will be developed to support the model’s use, as well as to share evidence on effective care intervention. HQO will evaluate the model’s ability to identify patients at-risk of future hospitalization, as well as the tool’s contribution to successful early care intervention.

Continuing analysis
While encouraging the use of HARP, and as new data sources become available, additional work may be undertaken to refine the model. A key priority is to incorporate predictive factors from primary care data. Data from electronic medical records, medical billings, and drug formularies could all contribute to strengthened risk prediction. As we look beyond acute data it may become increasingly difficult to identify factors that are of comparable predictive power, but identifying at-risk individuals prior to hospitalization would support earlier intervention. Primary care practices would have easy access to much of the data within their records systems, supporting the usability of HARP at that level of care.

Cost effective care
Finally, another benefit of predictive modeling is that groups can be analyzed on a cost basis. If we were to determine the average cost of caring for a person in the 90th percentile of risk, we can determine the maximum price of a cost-effective early intervention. These scenarios require certain estimations of avoided hospitalizations, but could provide a compelling case for investment in early patient care. This data would also enable cost-effectiveness evaluation by providing a benchmark for savings. We have identified this as an area of future work.

The statistical code needed to run the model is available, without charge, upon request. Please contact the Health System Research Team at hsr@cihi.ca for further information.
Acknowledgments

HQO and CIHI would like to thank those involved in the initiation of Hospital Admission Risk Prediction (HARP), the development of the model, and the writing of this report.

We would like to give particular acknowledgment to the contributions of the following experts for their guidance on model development and review of results:

- Dr. Walter P Wodchis (Chair), Associate Professor, University of Toronto
- Dr. Chaim Bell, General Internal Medicine, Mount Sinai Hospital, University Health Network
- Dr. Liisa Jaakkimainen, Department of Family and Community Medicine, University of Toronto
- Dr. David Frost, Division of General Internal Medicine, Toronto Western Hospital, University Health Network, University of Toronto
- Dr. Susan Law, Vice President, Academic Affairs, St. Mary’s Hospital Centre
- Dr. Samir Sinha, Director of Geriatrics, Mount Sinai Hospital, University Health Network
- Dr. Howard Abrams, Division of General Internal Medicine, Toronto Western Hospital, University Health Network, University of Toronto

The leadership team that provided guidance and support of the project included:

- Rebecca Comrie, Manager, Health Quality Ontario
- Gail Dobell, Director, Evaluation, Health Quality Ontario
- Kathleen Morris, Director, Health System Analysis and Emerging Issues, Canadian Institute for Health Information
- Jeremy Veillard, Vice President, Research and Analysis, Canadian Institute for Health Information

The project team responsible for the development of the model and report included:

- Allie Chen, Senior Analyst, Canadian Institute for Health Information
- Xi-Kuan Chen, Program Lead, Canadian Institute for Health Information
- Katerina Gapanenko, Manager, Canadian Institute for Health Information
- Ali Moses McKeag, Project Lead, Canadian Institute for Health Information
- Angus Steele (Project Lead), Senior Specialist, Health Quality Ontario

A special thanks to Dr. Geraint Lewis of the Nuffield Trust who provided advice early in the development of the project, and contributed helpful insight into the experience of predictive modelling in the United Kingdom. We would also like to acknowledge the contributions of the HQO Communications team, the bestPATH team, Erik Hellsten, and Sophia Ikura.
REFERENCES


