A Comparison of Methods for Forecasting Emergency Department Visits for Respiratory Illness Using Telehealth Ontario Calls

Alexander G. Perry, MSc,[1,2] Kieran M. Moore, MD,[1,3] Linda E. Levesque, PhD,[1,4] C. William L. Pickett, PhD,[3,4] Michael J. Korenberg, PhD[2]

ABSTRACT

Objectives: Anticipating increases in hospital emergency department (ED) visits for respiratory illness could help time interventions such as opening flu clinics to reduce surges in ED visits. Five different methods for estimating ED visits for respiratory illness from Telehealth Ontario calls are compared, including two non-linear modeling methods. Daily visit estimates up to 14 days in advance were made at the health unit level for all 36 Ontario health units.

Methods: Telehealth calls from June 1, 2004 to March 14, 2006 were included. Estimates generated by regression, Exponentially Weighted Moving Average (EWMA), Numerical Methods for Subspace State Space Identification (N4SID), Fast Orthogonal Search (FOS), and Parallel Cascade Identification (PCI) were compared to the actual number of ED visits for respiratory illness identified from the National Ambulatory Care Reporting System (NACRS) database. Model predictor variables included Telehealth Ontario calls and upcoming holidays/weekends. Models were fit using the first 304 days of data and prediction accuracy was measured over the remaining 348 days.

Results: Forecast accuracy was significantly better ($p<0.0001$) for the 12 Ontario health units with a population over 400,000 (75% of the Ontario population) than for smaller health units. Compared to regression, FOS produced better estimates ($p=0.03$) while there was no significant improvement for PCI-based estimates. FOS, PCI, EWMA and N4SID performed worse than regression over the remaining smaller health units.

Conclusion: Telehealth can be used to estimate ED visits for respiratory illness at the health unit level. Non-linear modeling methods produced better estimates than regression in larger health units.

Key words: Forecasting; surveillance; respiratory infections; mathematical model; hospital planning

La traduction du résumé se trouve à la fin de l'article.

Can J Public Health 2010;101(6):464-69.

In Canada and abroad, influenza and other respiratory illnesses place a significant burden on hospitals, potentially leading to emergency department (ED) overcrowding and bed shortages.[1-3] The Ontario Health Plan for an Influenza Pandemic recognizes the potential impact of surges due to influenza on acute care hospitals that already operate at, or near, capacity.[4] Strategies to mitigate the impact of surges could include opening alternative influenza assessment, treatment and referral centres.[4,5]

Timing and coordination of measures designed to reduce ED visit surges and thereby ensure access to care are important. There are potential cost and health impacts if they are applied too soon or too late.[4] The role of public health agencies in coordinating such measures has been recognized. In a Decision Document related to the H1N1 Pandemic, the Ontario Ministry of Health and Long-Term Care (MOHLTC) identified public health units as the lead coordinating agency for influenza pandemic planning locally, and specifically stated that "Primary care and hospitals are encouraged to engage with local public health units on the planning for and implementation of alternate assessment, treatment, and referral strategies…".[5] This recommendation is consistent with the Walker[6] and Campbell[7] reports commissioned after SARS stating that public health and hospitals must work together to coordinate responses to public health threats.

Although it has been suggested that surveillance information be used to estimate demand for acute care services to allow timely activation of measures to reduce hospital visit surges,[3,8] at present there is no method of measuring, in real-time, the number of hospital visits for respiratory illness across Ontario. Hospitals generally do not provide this information in real-time to a central database. The use of syndromic surveillance systems is beginning to change this by collecting real-time triage information on ED visits; however, there are logistical, cost and political barriers to collecting real-time hospital data across the province.

Because it provides province-wide coverage and is available electronically in near real-time from a single central database, Telehealth Ontario, a free 24-hour health advice telephone service offered throughout the province, is a good data source for influenza surveillance.[4,9] Previous research by van Dijk et al. has shown a significant correlation between calls to Telehealth for respiratory complaints and emergency visits for respiratory illness at the provincial level up to two weeks in advance.[9]

The objective of this study was to determine if Telehealth could be used to estimate the current and future number of ED visits for

Author Affiliations

1. Kingston, Frontenac and Lennox & Addington Public Health, Kingston, ON
2. Department of Electrical and Computer Engineering, Queen's University, Kingston, ON
3. Department of Emergency Medicine, Queen's University, Kingston, ON
4. Department of Community Health and Epidemiology, Queen's University, Kingston, ON

Correspondence: Alexander Perry, Kingston, Frontenac and Lennox & Addington Public Health, 221 Portsmouth Ave., Kingston, ON K7M 1V5, Tel 613-549-1232, ext. 1517, Fax: 613-549-7896, E-mail: alexander.perry@queensu.ca

Acknowledgements: This work was funded in part by an Ontario Graduate Scholarship (OGS) and Natural Sciences and Engineering Research Council of Canada (NSERC) scholarship held by A.G. Perry.

Conflict of Interest: None to declare.

respiratory illness at the health unit level for each of the 36 health units in Ontario. Although the earlier study by van Dijk indicates Telehealth is potentially useful for estimating ED visits at the provincial level, only simple cross-correlation was used to measure the strength of the association between calls and visits. No method is proposed to provide an estimate of visits from calls and the call-visit relationship is not examined at the health unit level. The study does not examine the possibility of non-linear relationship between calls and visits. To address this, the performance of five different forecasting methods that use calls to forecast ED visits is compared, including two techniques capable of capturing non-linear relationships between calls and visits. The methods examined include Exponentially Weighted Moving Average (EWMA), regression, subspace state space identification (N4SID), Fast Orthogonal Search (FOS), and Parallel Cascade Identification (PCI).

EWMA, regression, and Autoregressive Integrated Moving Average (ARIMA) have been applied in forecasting emergency department visits.[8,10,11] It can be shown that ARIMA models can be represented in state-space form[12,13] and in this study the N4SID method of identifying the parameters of a state-space model is used.[14] To investigate the possibility that non-linear relationships exist between calls and visits and are useful in estimating visits from calls, we employed two non-linear techniques for modeling time series: FOS and PCI.[15,16]

METHODS

Hospital ED visits and Telehealth Ontario calls from June 1, 2004 to March 14, 2006 (652 days) were included in the study. The study period was limited by availability of Telehealth data. Hospital ED visits for respiratory complaints were obtained from the Canadian Institute of Health Information (CIHI) National Ambulatory Care Reporting System (NACRS) database for the fiscal years 2004-2005 and 2005-2006. Telehealth Ontario calls were obtained from the MOHLTC. Information in each data set consisted of date of call or visit, patient age, and reason for visit or call. To protect the identity of individuals, all personal health information used in the study were de-identified and postal code information was limited to the forward sortation area (FSA). Ethics approval for the study was obtained from the Queen's University Health Sciences and Affiliated Teaching Hospitals Research Ethics Board.

Emergency visits for respiratory complaints were identified from the NACRS database using a set of International Classification of Disease Revision 10 Canadian Enhancement (ICD10-CA) codes. These codes, presented in Table 1, were obtained by translating ICD-9 codes found by previous research to be highly correlated with positive lab results for respiratory pathogens.[17] A similar set of codes had been used in the van Dijk study.[9] There were 548,438 visits for respiratory illness during the study period. Records with missing, invalid or out-of-province postal codes were excluded, leaving 535,185 visits in the analysis.

Each call to Telehealth Ontario is assigned to one of 486 call guidelines based on the reason for the call.[9] Calls for respiratory illness were identified using the same subset of these call guildelines as used by van Dijk.[9] There were 187,426 calls for respiratory complaints according to this set of guidelines over the study period. Of these, calls with missing, invalid or out-of-province FSA were excluded, leaving 177,601 calls in the analysis.

To construct time series of the daily number of Telehealth Ontario calls and daily number of emergency visits for respiratory illness for each health unit, a mapping between postal code FSA and health unit was constructed by linking dissemination areas using the Statistics Canada Health Region Boundary File[18] and Postal Code Conversion File.[19] Because an FSA may overlap several health units, the geographic regions used in the analysis approximate rather than exactly represent health unit regions. However, the same mapping was used for calls and visits meaning that calls and visits were aggregated over exactly the same mutually exclusive geographic regions.

Table 1. ICD10-CA Codes Used to Identify Hospital Visits for Respiratory Illness and Influenza

ICD10-CA Code	ICD1-CA Description
B34.9	Viral infection, unspecified
H66.9	Otitis media, unspecified
J00	Acute nasopharyngitis (common cold)
J01.9	Acute sinusitis, unspecified
J06.8	Other acute upper respiratory infections of multiple sites
J39.9	Disease of upper respiratory tract, unspecified
J06.9	Acute upper respiratory infection, unspecified
J20.0	Acute bronchitis due to mycoplasma pneumoniae
J20.1	Acute bronchitis due to haemophilus influenzae
J20.2	Acute bronchitis due to streptococcus
J20.3	Acute bronchitis due to coxsackievirus
J20.4	Acute bronchitis due to parainfluenza virus
J20.5	Acute bronchitis due to respiratory syncytial virus
J20.6	Acute bronchitis due to rhinovirus
J20.7	Acute bronchitis due to echovirus
J20.8	Acute bronchitis due to other specified organisms
J20.9	Acute bronchitis, unspecified
J18.8	Other pneumonia, organism unspecified
J18.9	Pneumonia, unspecified
J10.0	Influenza with pneumonia, influenza virus identified
J11.0	Influenza with pneumonia, virus not identified
J10.1	Influenza with other respiratory manifestations, influenza virus identified
J11.1	Influenza with other respiratory manifestations,virus not identified
J10.8	Influenza with other manifestations, influenza virus identified
J11.8	Influenza with other manifestations, virus not identified
J40	Bronchitis, not specified as acute or chronic
R50.0	Fever with chills
R50.1	Persistent fever
R50.9	Fever, unspecified
R05	Cough

Forecasts of the daily number of ED visits for respiratory illness at the health unit level were generated by each of the five different modeling methods for each of the 36 health units in Ontario. Models were constructed to produce estimates of the daily number of visits, $\hat{y}$, for given health unit q days ahead, where q ranged from 0 to 14. A maximum forecasting horizon of 14 days was chosen because previous research indicated that calls were significantly correlated with visits up to two weeks in advance.[9] Each method had as inputs the daily aggregate number Telehealth calls, x_c, for the health unit being considered and an indicator variable for upcoming holidays and weekends, x_h (except the EWMA model which used only the former). Table 2 presents details of the models used. The variable n indicates the time index in days.

All models were implemented in MATLAB.[21] Models were fit using the first 304 days of Telehealth data (training data set) and performance was assessed on the remaining data (validation data set). A total of 2,700 models were created (5 methods, 15 lead times, 36 health units). Forecasting accuracy was assessed using the mean square error, %MSE, defined as:

$$\%MSE = \frac{\overline{\left(y(n)-\hat{y}(n)\right)^2}}{\overline{\left(y(n)-\overline{y(n)}\right)^2}}$$

where y(n) is the actual number of emergency visits at day n, $\hat{y}(n)$ is the estimated number of ED visits at day n, and the overbar "$^{-}$"

Table 2. Forecasting Models and Implementation Details

Model and Identification Method	Parameter Definitions and Implementation Details
EWMA Model $\hat{y}(n+q)=\alpha\hat{y}(n+q-1)+\beta x_c(n)$ Model coefficients identified by least-squares fit	α and β are the model coefficients. Note that the models did not use the actual number of visits on previous days, y(n+q-1) as a predictor of future visits. This assumption was made as this information is not currently available in Ontario – hospitals are only required to report visit information to CIHI at the end of the fiscal year. Instead the previous model estimate was used.
Regression Model $\hat{y}(n+q)=\beta_0+\sum_{i=0}^{L}\beta_{i+1}x_c(n-i)+\beta_{L+1}x_h(n+q)$ Model coefficients identified by least-squares fit	β_i are the model coefficients. L is the maximum lag in the calls allowed in the model and was set to 10.
State Space Model $S(n+1)=AS(n)+BX(n)$ $\hat{y}(n+q)=CS(n)+DX(n)$ Model identified by N4SID algorithm[14] implemented in the MATLAB System Identification Toolbox[20,21]	X(n) is a vector of calls, x_c, at time lags up to 10 and holidays and upcoming holidays, x_h. S(n) is a vector of system states. Three state variables were allowed in the model. A, B, C, D are matrices of appropriate dimensions.
Parallel Cascade Model $\hat{y}(n+q)=\sum_{j=1}^{C}\left[\sum_{i=1}^{P}a_{ij}\left(\sum_{k=0}^{L}h_j(k)x_{R_j}(n-k)\right)^i\right]$ Model identified using Parallel Cascade Identification (PCI)[16]	$R_i\in\{c,h\}$, indicating that either calls or holidays could be an input to a given cascade. In the case of Ri=h, k can take negative values up to -q, otherwise k≥0. C is the number of cascades and was set to 3. P is the order of the polynomial describing the static non-linearity and was set to 2. L is the memory length of the dynamic linear element and was set to 10. $h_j(k)$ is the impulse response of the dynamic linear element of the j^{th} cascade. Up to a third order cross-correlation was allowed when determining the impulse response of the dynamic linear element in the PCI model.
Non-linear Difference Equation Model $\hat{y}(n+q)=\sum_{i=1}^{M}a_i\prod_{j=1}^{C_i}x_c(n-t_{c,i,j})\prod_{k=1}^{H_{ii}}x_h\left(n-t_{k,i,k}\right)$ Model identified using Fast Orthogonal Search (FOS)[15]	M is the number of terms in the model and was set to 8. a_i is the coefficient of the i^{th} term in the model. C_j is the number of x_c factors in the i^{th} term; $0\leq C_j\leq 3$ for any given term. $t_{c,i,j}$ is the lag of the j^{th} x_c factor in the i^{th} term; $0\leq t_{c,i,j}\leq 10$ for any given term. H_i is the number of x_h factors in the i^{th} term; $0\leq H_k\leq 2$ for any given term. $t_{h,i,j}$ is the lag of the j^{th} x_h factor in the i^{th} term; $0\leq t_{h,i,j}\leq 10$ for any given term.

Table 3. Population,[22] Call-to-visits Ratio, and %MSE for Four-day-ahead Predictions

Health Unit	2005 Population		Call-to-Visits Ratio	%MSE (4 days ahead)				
	n	(%)*		Regression	EWMA	FOS	PCI	State Space
City of Toronto Health Unit	2,627,821	(20.9)	0.64	38.0	58.3	36.0	38.6	44.1
Peel Regional Health Unit	1,217,457	(9.7)	0.70	45.8	63.9	43.1	45.5	49.7
York Regional Health Unit	917,813	(7.3)	0.76	52.5	79.2	52.1	57.5	59.8
City of Ottawa Health Unit	836,451	(6.7)	0.64	46.1	59.7	43.5	48.5	50.7
Durham Regional Health Unit	574,757	(4.6)	0.50	58.6	89.7	54.0	57.1	60.8
City of Hamilton Health Unit	520,242	(4.1)	0.27	58.9	59.5	52.8	47.4	53.4
Waterloo Health Unit	485,414	(3.9)	0.58	61.7	85.5	66.3	58.6	71.1
Simcoe Muskoka District Health Unit	484,095	(3.9)	0.27	51.9	68.0	51.0	47.1	47.1
Halton Regional Health Unit	443,732	(3.5)	0.67	56.8	104.0	61.4	60.5	64.1
Middlesex London Health Unit	434,696	(3.5)	0.35	60.0	85.8	69.2	65.4	65.8
Niagara Regional Area Health Unit	434,409	(3.5)	0.21	49.4	55.4	55.7	48.0	55.4
Windsor Essex County Health Unit	404,665	(3.2)	0.38	117.4	90.6	106.1	89.9	98.2

* Percentage of the total estimated 2005 Ontario population: 12,565,446.[22]

denotes the time average. A multi-level regression model (with regression as referent), accounting for correlation in the %MSE within health units, was used to assess the effects of forecasting method, health unit population, and forecasting lead time on the %MSE across all models and all 36 health units. SAS was used to fit this model. The interaction between method and population was examined.

RESULTS

Table 3 presents a description of Ontario health units with a population greater than 400,000 in 2005.[22] Collectively, these health units represent about 75% of the Ontario population. The %MSE for each of the methods for four-day-ahead predictions is presented in the table, in addition to the ratio of the median number of daily Telehealth calls to the median number of daily ED visits. The call-to-visits ratio was found to be associated with model accuracy.

To give the reader a sense of the meaning of the %MSE values reported in Table 3, Figure 1 shows a plot comparing the actual to four-day-ahead estimated number of ED visits for the FOS, regression, and PCI models over part of the validation data for the city of Toronto Health Unit (EWMA and State Space are omitted for clarity). Figure 2 shows the predicted number of visits over the next one-week period (dashed lines) compared with the actual ED visits (solid line) that occurred during that period. The one-week-ahead weekly predictions were obtained by summing the one-through seven-day-ahead predictions. Aggregating over a seven-day period removes the well-known weekly cyclical pattern in visits[23] and more clearly shows ability to predict beyond these known trends.

In health units with populations of more than 400,000, the multi-level regression model indicated that FOS offered better

Figure 1. Four-day-ahead emergency visit forecasts for the City of Toronto Health Unit

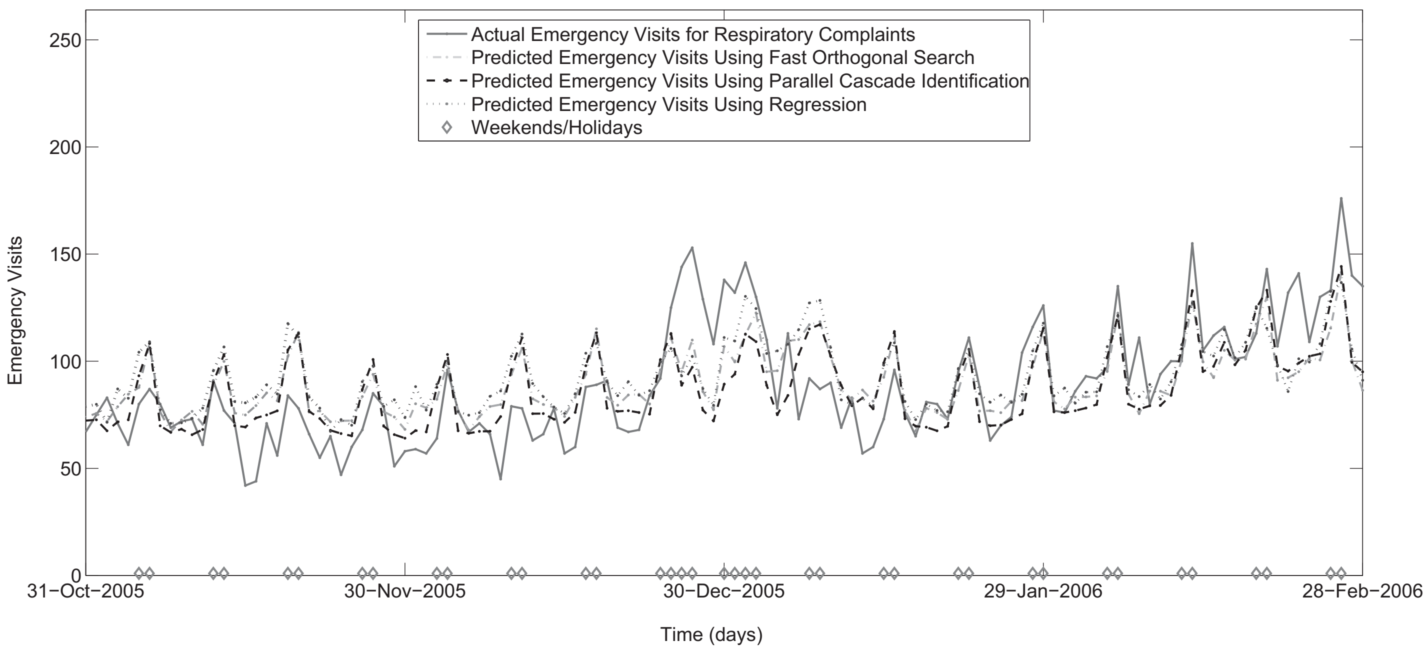
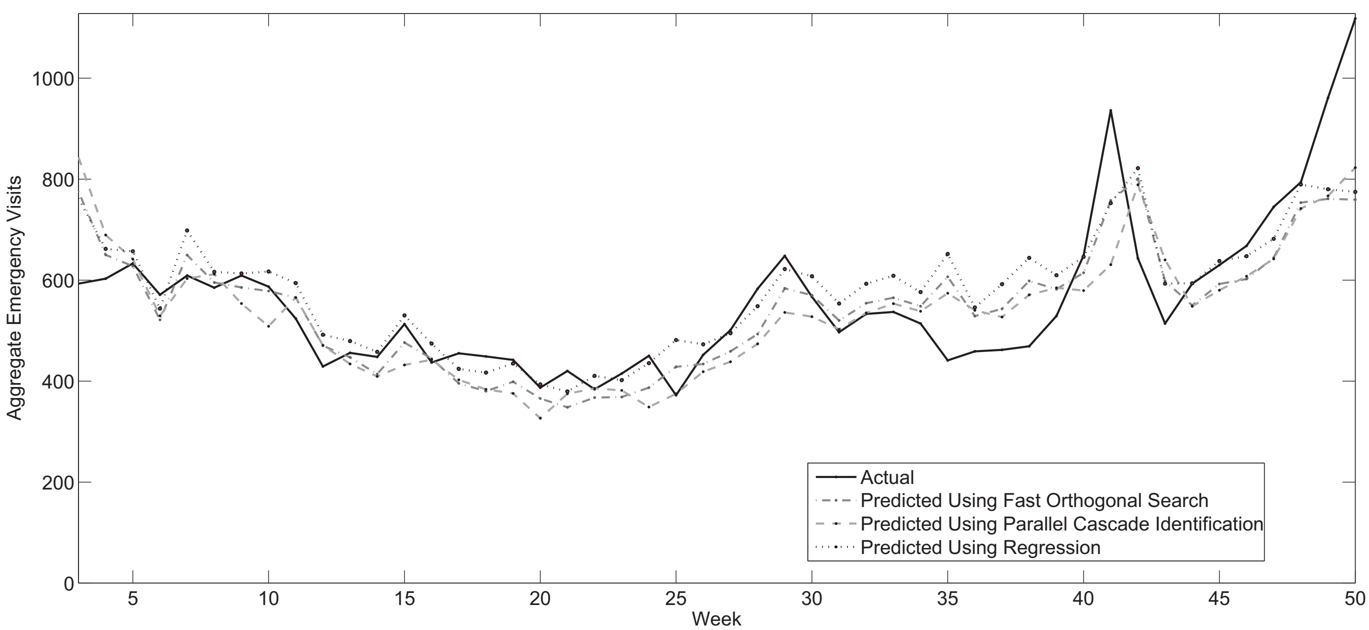


Figure 2. One-week-ahead forecasts for the City of Toronto Health Unit



performance (p=0.03) than regression. PCI did not provide statistically significant better performance over regression (p=0.44) while both the EWMA and state-space method provided worse performance (p<0.0001). The %MSE in forecast accuracy was lower for health units with a population of more than 400,000 (p<0.0001). The interaction terms between forecasting method and population size for FOS and PCI were statistically significant (p<0.001), indicating that these methods provided poorer performance than regression for health units with populations less than 400,000 while interactions for the EWMA and state-space/N4SID models were not significant (p=0.24 for each). Forecast accuracy decreased with lead time (p<0.0001).

DISCUSSION

The results of this study are consistent with those obtained by van Dijk et al. where Telehealth was shown to be correlated with ED visits at the provincial level up to two weeks in advance.[9] The results indicate that Telehealth can be used to estimate future visits at the health unit level. Estimates are better for health units with larger populations and non-linear modeling methods produced more accurate estimates for these health units. Because no knowledge of the actual number of past visits is used in any model over the validation data, the statistically significant decrease in prediction accuracy with increasing forecasting lead time provides evidence that there is information about the future number of visits in the call time series. Suppose the call time series were not a predictor of ED visits, and for example only the mean number of visits over the training data and information about upcoming holidays and weekends were useful in creating forecasts. Then we would expect a non-significant change in forecasting error with lead time as upcoming holidays/weekends are always precisely known regardless of the forecast horizon.

A major limitation of this study was that only two years of Telehealth data were available. Due to this limitation and because only one flu season was available to test prediction performance, it is difficult to draw firm conclusions about the performance of the various methods or the importance of considering nonlinearity. The use of least-squares to fit coefficients for the models may introduce some bias as the distribution of visit counts may be more Poisson than Gaussian and this may have degraded predictive performance. However, the error introduced by this assumption is likely negligible compared to the error due to unexplained variation. An important assumption made in developing the forecasting models was that the Telehealth calls–ED visits relationship does not change over time: we built a predictive model using approximately the first year of data and then tested it over the second year of data. However, since calling behaviours could change over time, for example in response to factors such as season and promotion of the Telehealth service, this assumption could have been violated. Since data confirming past visits is not available in real-time, models cannot be updated on an ongoing basis. Collection of real-time data through syndromic surveillance systems may address this limitation in the future. It was not possible to account for individuals who made calls in one area but sought care in another. This might have influenced results where geographic boundaries separate a large population from the centre of care used by this population.

The fact that accuracy of ED visit prediction using Telehealth is better for larger health units might be explained by the fact that call-to-visits ratios are much higher for health units with larger populations (correlation coefficient 0.70, $p<0.001$). The call-to-visits ratio is simply a measure of the number of calls, adjusted for the number of visits. It might be expected that more calls provide more information that allows the number of visits to be predicted. The fact that non-linear methods show improved performance relative to regression only over health units with larger population might also be explained considering the call-to-visits ratio as non-linear methods may provide benefit only when richer call data are available.

Since the first terms selected by FOS included three to four linear terms of the eight terms in the model, it is not surprising that the non-linear methods provided only limited improvements in prediction accuracy. The FOS method works by adding terms one at a time to the model from a pool of candidates, consisting of a set of all possible predictors and their powers and cross-products, on the basis of how much they reduce error in the model accuracy over the training data. An advantage of the FOS method over normal regression is that it can search for terms in the model including power and cross-product (interaction) predictor terms much more quickly than forward, backward, or step-wise selection with regression. The number of candidate predictor variables can be large in time series analysis when the influence of these predictors at multiple lags is considered. For example, the FOS algorithm we used was able to search 28,392 candidates to generate an eight-term model in approximately 17 seconds using a standard laptop computer. It should be noted that PCI and FOS performance was better than regression for health units with populations of more than 400,000 ($p<0.0001$) over both the training and the validation data. Model performance can degrade if over-fitting occurs on the training data, and therefore performance should be examined over both training and validation data when comparing methods.

It is not surprising that the EWMA and State-Space methods perform worse than the regression, FOS and PCI methods since these methods explicitly (EWMA) or implicitly (state-space – through the use of state variables) use value of past estimates to produce future estimates. Inaccuracies in past predictions can accumulate to impact future predictions. Performance of these methods might be improved should the actual number of past ED visits be available. In this case, FOS could also be augmented to include terms with past ED visits. Finally, the practical difference in the predictions made using different methods should be considered in addition to the statistical significance of the difference in their accuracy in the context of the intended application.

REFERENCES

1. Menec V, Bruce S, MacWilliam L. Exploring reasons for bed pressures in Winnipeg acute care hospitals. *Can J Aging* 2005;24(Supplement 1):121-31.
2. Menec V, Black C, MacWilliam L, Aoki F. The impact of influenza-associated respiratory illnesses on hospitalizations, physician visits, emergency room visits, and mortality. *Can J Public Health* 2003;94(1):59-63.
3. Hanratty B, Robinson M. Coping with winter bed crises. *BMJ* 1999;319:1511-12.
4. Ontario Ministry of Health and Long-Term Care. Ontario Health Plan for an Influenza Pandemic, 2008.
5. Ontario Ministry of Health and Long-Term Care. Pandemic H1N1 (pH1N1) Alternate Influenza Assessment, Treatment, and Referral Services. Decision Document. August 2009.
6. Walker D, Keon W, Laupacis A, Low D, Moore K, Kitts J, et al. The Ontario Expert Panel on SARS and Infectious Disease Control: For the Public's Health: A Plan of Action—Final Report of the Ontario Expert Panel on SARS and Infectious Disease Control, 2004.
7. Campbell A. The SARS Commission Interim Report: SARS and Public Health in Ontario. Ontario Ministry of Health and Long-Term Care, April 2004.
8. Jones S, Joy M. Forecasting demand of emergency care. *Health Care Management Sci* 2002;5:297-305.
9. van Dijk A, McGuinness D, Rolland E, Moore KM. Can Telehealth Ontario respiratory call volume be used as a proxy for emergency department respiratory visit surveillance by public health? *CJEM* 2008;10(1):18-24.
10. Brillman JC, Burr T, Forslund D, Joyce E, Picard R, Umlan E. Modeling emergency department visit patterns for infectious disease complaints: Results and application to disease surveillance. *BMC Med Informatics Decision Making* 2005;5(4).
11. Reis B, Mandl K. Time series modeling for syndromic surveillance. *BMC Med Informatics Decision Making* 2003;3(2).
12. Gilbert P. State space and ARMA models: An overview of the equivalence. Bank of Canada Working Paper 93-4, 1993.
13. Makridakis S, Wheelwright S, Hyndman R. *Forecasting: Methods and Applications,* 3rd ed. New York, NY: Wiley, 1998.
14. Overschee P, DeMoor B. N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica* 1994;30(1):75-93.
15. Korenberg M. A robust orthogonal algorithm for system identification and time-series analysis. *Biological Cybernetics* 1989;60:267-76.
16. Korenberg M. Parallel cascade identification and kernel estimation for nonlinear systems. *Ann Biomedical Engineering* 1991;19:429-55.
17. Marsden-Haug N, Foster V, Gould P, Elbert E, Wang H, Pavlin J. Code-based syndromic surveillance for influenza-like illness by international classification of diseases, Ninth Revision. *Emerg Infect Dis* 2007;13(2):207-16.
18. Statistics Canada. Health Regions: Boundaries and Correspondence with Census Geography. 82-402-XIE, 2007.
19. Statistics Canada. Postal Code Conversion File (PCCF), Reference Guide. 92F0153GIE, 2006.
20. Ljung L. *MATLAB System Identification Toolbox User's Guide,* Version 6 ed. Natick, MA: The MathWorks, 2003.
21. The MathWorks Inc. MATLAB Version 7 Release 14. 2003.
22. Statistics Canada. Table 109-5315 - Estimates of population (Census and administrative data), by age group and sex, Canada, provinces, territories, health regions and peer groups, annual (number) (table), CANSIM (database), Using E-STAT (distributor). Available at: http://estat.statcan.gc.ca/cgi-

win/cnsmcgi.exe?Lang=E&EST-Fi=EStat/English/CII_1-eng.htm (Accessed August 10, 2009).
23. Chan B, Schull MJ, Schultz SE. Atlas Report: Emergency Department Services in Ontario 1993-2000. Toronto, ON: Institute for Clinical Evaluative Sciences, 2001.

Received: November 12, 2009
Accepted: July 23, 2010

RÉSUMÉ

Objectifs : Si l'on pouvait prévoir les augmentations des visites à l'urgence associées aux maladies respiratoires, on réduirait leur impact sur les hôpitaux en ciblant mieux, par exemple, les ouvertures de cliniques secondaires. Nous avons comparé cinq méthodes d'estimation du nombre de visites à l'urgence associées aux maladies respiratoires d'après les appels à Télésanté Ontario, y compris deux méthodes de modélisation non linéaires. Nous avons estimé le nombre de visites quotidiennes jusqu'à 14 jours à l'avance pour chacune des 36 circonscriptions sanitaires de l'Ontario.

Méthode : Nous avons inclus les appels reçus par Télésanté entre le 1er juin 2004 et le 14 mars 2006. Les estimations produites par la régression multivariée, la moyenne mobile à pondération exponentielle (MMPE), les méthodes numériques d'identification des sous-espaces (N4SID), la recherche orthogonale rapide (ROR) et l'identification parallèle en cascade (IPC) ont été comparées au nombre réel de visites à l'urgence associées aux maladies respiratoires enregistré dans la banque de données du Système national d'information sur les soins ambulatoires (SNISA). Les variables prédictives des modèles étaient les appels à Télésanté, les jours fériés à venir et les fins de semaine. Les modèles ont été ajustés selon les 304 premiers jours, et la précision des prédictions a été mesurée au cours des 348 jours suivants.

Résultats : La précision des prévisions était significativement supérieure ($p<0,0001$) dans les 12 circonscriptions sanitaires de plus de 400 000 habitants (75 % de la population de l'Ontario) que dans les circonscriptions plus petites. La ROR a produit les meilleures estimations ($p=0,03$), tandis que l'IPC n'apportait aucune amélioration significative. Les méthodes ROR, IPC, MMPE et N4SID ont produit de moins bons résultats que la régression dans les petites circonscriptions sanitaires.

Conclusion : Télésanté Ontario peut être utilisée pour estimer les visites à l'urgence associées aux maladies respiratoires dans les circonscriptions sanitaires. Les méthodes de modélisation non linéaires produisent de meilleures estimations que la régression dans les circonscriptions qui englobent la majorité de la population.

Mots clés : prévision; surveillance; infections de l'appareil respiratoire; modélisation mathématique; planification hôpitaux