
Identifying cases of congestive heart failure from administrative data: a validation study using primary care patient records

S. E. Schultz, MA, MSc (1); D. M. Rothwell, MSc (2); Z. Chen, MD (1); K. Tu, MD (1, 3, 4)

This article has been peer reviewed.

Abstract

Introduction: To determine if using a combination of hospital administrative data and ambulatory care physician billings can accurately identify patients with congestive heart failure (CHF), we tested 9 algorithms for identifying individuals with CHF from administrative data.

Methods: The validation cohort against which the 9 algorithms were tested combined data from a random sample of adult patients from EMRALD, an electronic medical record database of primary care physicians in Ontario, Canada, and data collected in 2004/05 from a random sample of primary care patients for a study of hypertension. Algorithms were evaluated on sensitivity, specificity, positive predictive value, area under the curve on the ROC graph and the combination of likelihood ratio positive and negative.

Results: We found that that one hospital record or one physician billing followed by a second record from either source within one year had the best result, with a sensitivity of 84.8% and a specificity of 97.0%.

Conclusion: Population prevalence of CHF can be accurately measured using combined administrative data from hospitalization and ambulatory care.

Keywords: congestive heart failure, validation studies, epidemiologic methods, population prevalence

Introduction

Hospital discharge abstracts^{1,2} have traditionally been used to identify those patients with congestive heart failure (CHF) who present to hospital or who are hospitalized for other conditions but have CHF listed as a co-morbidity. In their recent systematic review of validation studies of algorithms to identify CHF from administrative data, Saczynski et al.³ found this to be true for 25 of 35 studies listed.

Compared with hospital records, the use of hospital discharge abstracts to identify patients with CHF has been found to be highly accurate.^{4,5} However, with improvements in treatment and decreases in hospital resources, more patients with heart failure are being successfully managed in the community. As a result, they may never show up in the hospital discharge data or else not until their disease is in the advanced stages. Thus, using hospital data alone will probably underestimate the incidence and prevalence of CHF.

Validated algorithms using combinations of physician billing data and hospital discharge abstracts have been developed to identify patients with chronic disease conditions that do not necessarily require hospitalization, for example, hypertension, diabetes, ischemic heart disease and asthma.⁶⁻⁹ However, of the 35 studies listed in the systematic review conducted by Saczynski et al.,³ only 9 used data from both hospital discharges and ambulatory claims data, and only 2 were also population-based, although the population was still limited to patients enrolled in a large managed-care organization.^{10,11}

The purpose of our study was to determine the most suitable algorithm of administrative data to identify patients with CHF in Ontario, Canada. We used information within primary care physician outpatient electronic medical records (EMRs) and fee-for-service primary care physician charts to assess the validity and reliability of various combinations of physician billing data and hospital discharge data.

Methods

Data sources

Validation cohort

The validation cohort used in this study comprised data from two sources. The first was collected through the Canadian Cardiovascular Outcomes Research Team (CCORT) from 17 physicians using Practice Solutions® Electronic Medical

Author references:

1. Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada
2. Ottawa Hospital Research Institute, Ottawa, Ontario, Canada
3. Department of Family and Community Medicine – University of Toronto, Toronto, Ontario, Canada
4. Toronto Western Hospital Family Health Team – University Health Network, Toronto, Ontario, Canada

Correspondence: Susan E. Schultz, Institute for Clinical Evaluative Sciences G1-06, 2075 Bayview Ave., Toronto, ON M4N 3M5; Tel.: 416-480-4055 ext. 1-3788; Fax: 416-480-6048; Email: sue.schultz@ices.on.ca

Records (EMR) that contributed their patient records into the Electronic Medical Record Administrative data Linked Database (EMRALD). Physicians participating in this study had to have been using the EMR for a minimum of 2 years in order to have an EMR populated with a full practice of patients. Data from the EMR were extracted from June to December of 2007, anonymized, encrypted and then transferred electronically to the Institute for Clinical Evaluative Sciences (ICES) in a secure fashion. ICES is a prescribed entity under the Ontario Personal Health Information Protection Act, which means the organization can receive and use health information, without consent, for analysis and to compile statistical information about the Ontario health care system. Data can be collected from a variety of sources, including the Ministry of Health and Long-Term Care, hospitals and physicians, provided ICES has in place policies, practices and procedures that have been audited and approved by the Information and Privacy Commissioner of Ontario. The data used in this study were handled as per ICES' standard operating procedures to preserve patient privacy and confidentiality.

The total eligible EMR patient population consisted of 19 376 active adult patients aged 20 years or more. "Active" was defined as rostered to the participating physician, having at least 2 visits in the last 3 years and a valid Ontario Health Insurance Plan (OHIP) card. Data were abstracted from a 5% random sample of patient charts ($n=969$) by three trained abstractors. Inter-observer reliability calculated on global agreement on the presence or absence of CHF was very good (kappa score [κ] > 0.80).

The second data source was a random sample of patient charts abstracted from 76 fee-for-service family physician practices between December 2004 and August 2005 for validation of an administrative data-based algorithm to detect cases of hypertension.⁶ In this study, charts were abstracted from a random sample of all eligible patients ($n=2472$). Eligible patients were aged 38 years or more, regular and current patients of the practice with at least 2 visits during the previous 3

years and an OHIP health card number. Two abstractors abstracted the charts, once again with very good inter-observer reliability, and again calculated on the overall status of the presence or absence of CHF ($\kappa > 0.80$).

We identified patients with CHF from both data sources in a similar manner: the trained chart abstractors reviewed all available entries in the EMR or in the previous three years in the patient charts. Each entry was scored as to whether it indicated "definite" CHF, "possible" CHF, a family history of CHF or no CHF. A "definite" CHF entry meant there was explicit physician documentation stating that the patient had CHF or one or more of the following synonyms: biventricular failure, cardiac decompensation, cardiac failure, heart failure (right or left), pump failure, ventricular failure or wet lungs. Diagnostic test findings consistent with CHF were classified as "possible" CHF. If there was no mention of CHF or any of its synonyms, patients were considered as having no CHF. The resulting abstraction classification for each patient was tabulated and patients were considered to have CHF only if one or more entries denoted definite CHF.

To ensure consistency between the two data sources, only individuals from the EMR cohort aged 38 years or more were included. We did not expect this to introduce any bias as the EMR sample was a simple random sample from each physician practice. To enable analysis of these data, each patient's health card number was encrypted, yielding a unique identification number that could be linked to the Ontario administrative data holdings housed at ICES.

Administrative data sources

The administrative data sources used to detect cases of CHF in the population included the hospital discharge abstract database (DAD) and the same-day surgery database (SDS), maintained by the Canadian Institute for Health Information (CIHI), and the OHIP database of physician fee-for-service billings or shadow-billings. The DAD and SDS classify pre-2002 diagnoses using codes from the *International Classification of Diseases*,

9th Revision (ICD-9) and later ones using the 10th Revision (ICD-10); the OHIP database uses a modified version of ICD-8. The OHIP physician billing data records over 95% of office-based primary care physician encounters for Ontario residents. A diagnostic code for CHF in the OHIP physician billing database or in the CIHI hospitalization databases (DAD or SDS) was the most responsible diagnosis; otherwise, a co-morbid condition was taken as positively indicating CHF.

Diagnostic codes used to define CHF

The diagnostic codes used to define CHF vary considerably. In the ICD-9 schema, CHF is most often defined as ICD-9 428. Lee et al.⁴ validated ICD-9 428 against two sets of clinical criteria using information from patients' hospital records and found it to be highly predictive. In turn, Vermeulen et al.¹³ compared the performance of ICD-9 428 and ICD-10 I50 and found them to be comparable. These two results address the question, "Do ICD-9 428 or ICD-10 I50 correctly identify cases of heart failure?" However, they do not speak to the issue of whether these codes are sufficient to detect *all* cases of CHF. Studies from other jurisdictions^{5,14,15} have used a broader range of diagnostic codes to identify cases of CHF from administrative data. We decided to compare two CHF definitions: a narrow definition that uses only ICD-9 428 and ICD-10 I500, I501 and I509, and a broader definition that also includes the codes for cardiomyopathy (ICD-9 425; ICD-10 I42) and pulmonary edema (ICD-9 514, 518.4; ICD-10 J81).

Algorithms tested

We tested 9 algorithms, which varied according to the data sources used and the length of time of follow-up. The performance of the various administrative data algorithms was evaluated against the manually abstracted CHF status from the patient's chart/EMR. These algorithms can be broadly divided into 3 groups. The first, algorithms 1 to 3, require only one record for a diagnosis of CHF but test the use of different data sources. Algorithms 4 to 6 require either one inpatient record or one ambulatory care record plus an additional record from

either source within a specific time period that varies between 1, 2 or 3 years. The third group, algorithms 7 to 9, are similar to 4 to 6 but use only ambulatory care data and require two ambulatory care records within 1, 2 or 3 years of follow-up (see Table 1).

We evaluated each algorithm with respect to its sensitivity, specificity and positive predictive value (PPV) and calculated 95% confidence intervals (CIs) using the binomial approximation method. All analyses were conducted using SAS version 9.2 (SAS Institute, Cary, NC, US).

We also estimated two summary measures, the receiver operating characteristics (ROC) curve and the likelihood ratio positive and negative (LR+ and LR-) to provide additional diagnostics. The ROC curve, originally developed to evaluate signal detection by radar operators, plots sensitivity (true positives) against 1 minus specificity (false positives). The closer the area under the curve (AUC) is to 1.00, the better the test.

The likelihood ratio graph plots the LR+, which is the ratio of sensitivity (the true positive rate) to 1 minus specificity (the false positive rate), against the LR-, which is the ratio of the false negative rate (1 - sensitivity) to the true negative rate (specificity).¹⁶ The LR+ measures the ability of the test to include those who have the condition while the LR- measures the ability to rule out those without the condition. Using the cut points first suggested by Jaeschke et al.,¹⁷ tests with

an LR+ greater than 10 and an LR- less than 0.1 are considered very useful, those with an LR+ between 5 and 10 and an LR- between 0.1 and 0.2 are considered moderately useful and those with an LR+ between 2 and 5 and an LR- between 0.2 and 0.5 are only somewhat useful. Tests with an LR+ less than 2 and an LR- more than 0.5 are of no use.

Results

The combined validation cohort comprised 2338 patients, with 99 definite for CHF according to their chart or EMR and 2239 without the condition (prevalence = 4.2%).

The age and gender distribution of patients in the validation cohort was similar to that of the 2006 Ontario population¹⁸ aged 38 years and older with a slight over-representation of women (56% in the cohort compared with 52% in the general population) and of individuals aged 65 years and older (32% in the cohort as opposed to 26% in the Ontario population aged 38 years and older). This was to be expected as our validation cohort was composed of individuals who make regular visits to a family physician, and both women and seniors are known to be more likely to visit a physician.¹⁹ The average age of our active adult cohort was 57.9 years, slightly higher than the average age (56.1 years) of Ontarian adults aged 38 years or older.

Taking a broad look across the three groups of algorithms (see Table 2), two

things become evident. The first is that follow-up time makes very little difference. The results for algorithms 4, 5 and 6 are nearly identical, and there is also very little difference in the performance of 7, 8 and 9. What appears to make more of a difference is the choice of data sources. For example, the sensitivity of algorithms 4, 5 or 6, which use both hospitalization and ambulatory care data, is at least 10 percentage points higher than that of algorithms 7, 8 or 9, which use only ambulatory care data, and 20 percentage points higher than hospital data alone.

The impact of data source on algorithm performance becomes particularly evident in the results for algorithms 1, 2 and 3. Algorithm 3, which requires only one record from any source for a diagnosis of CHF, had the highest sensitivity at 89.9% but the lowest specificity (93.5%) and the poorest PPV (38.0%). Inpatient data alone (algorithm 1) was the least sensitive, detecting only 60.6% of CHF cases, but it also had the highest specificity at 98.6% and the highest PPV (65.9%; Table 2). Negative predictive value (data not shown) was uniformly high, ranging from 99.6% for algorithm 3 to 98.2% for algorithm 1.

Turning to the results of the summary measures, it is interesting to note that the most useful algorithms are again those that use both ambulatory and hospitalization data. All the algorithms that use only one type of data, be it hospital or ambulatory care, are lower on the ROC curve or in the “somewhat useful” area of the LR graph (Figure 1). Looking at the ROC curve (Figure 2), the performances of algorithm 3 followed by 4, 5 and 6 are the best of the nine. The AUC for algorithm 3 was 0.917 and for 4, 5 and 6 was 0.909. When comparing the LR results (Figure 2), algorithms 3, 4, 5, 6 and 2 all fall within the “moderately useful” section of the graph, with algorithm 3 closest to the “very useful” section and algorithm 2 the furthest.

Selecting the best algorithm

The final choice for the best algorithm to identify CHF cases is between algorithms 3 and 4 (5 and 6 having been dropped from

TABLE 1
Administrative data algorithms tested against manually abstracted CHF status from the patient’s chart or EMR

Algorithm	Description
1	1 hospital record
2	1 ambulatory care record
3	1 hospital or ambulatory care record
4	1 hospital record alone OR 1 ambulatory care record followed by another record in 1 year
5	1 hospital record alone OR 1 ambulatory care record followed by another record in 2 years
6	1 hospital record alone or 1 ambulatory care record followed by another record in 3 years
7	2 ambulatory care records in 1 year
8	2 ambulatory care records in 2 years
9	2 ambulatory care records in 3 years

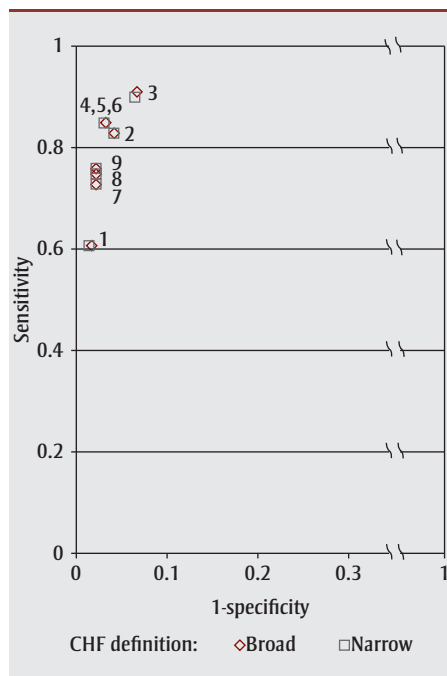
Abbreviations: EMR, electronic medical record; CHF, congestive heart failure.

TABLE 2
Sensitivity, specificity and PPV of nine different algorithms used to estimate CHF status from manually abstracted primary care physician data and administrative data

Algorithm number	Description	CHF definition	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)
1	1 CIHI record	Narrow	60.6 (50.8, 70.4)	98.6 (98.1, 99.1)	65.9 (56.0, 75.9)
		Broad	60.6 (50.8, 70.4)	98.3 (97.8, 98.8)	61.2 (51.4, 71.0)
2	1 OHIP claim	Narrow	82.8 (75.3, 90.4)	95.8 (94.9, 96.6)	46.3 (51.4, 71.0)
		Broad	82.8 (75.3, 90.4)	95.8 (94.9, 96.6)	46.3 (38.9, 53.7)
3	1 CIHI or OHIP claim	Narrow	89.9 (83.9, 95.9)	93.5 (92.5, 94.5)	38.0 (31.8, 44.3)
		Broad	90.9 (85.1, 96.7)	93.3 (92.3, 94.4)	37.7 (31.5, 43.8)
4	1 CIHI or 1 OHIP + 2nd claim (any source) in 1 year	Narrow	84.8 (77.7, 92.0)	97.0 (96.3, 97.7)	55.6 (47.6, 63.6)
		Broad	84.8 (77.7, 92.0)	96.8 (96.1, 97.5)	53.8 (45.9, 61.8)
5	1 CIHI or 1 OHIP + 2nd claim (any source) in 2 years	Narrow	84.8 (77.7, 92.0)	97.0 (96.1, 97.5)	55.3 (47.3, 63.3)
		Broad	84.8 (77.7, 92.0)	96.7 (96.0, 97.5)	53.5 (45.6, 61.4)
6	1 CIHI or 1 OHIP + 2nd claim (any source) in 3 years	Narrow	84.8 (77.7, 92.0)	96.9 (96.2, 97.6)	54.9 (46.9, 62.9)
		Broad	84.8 (77.7, 92.0)	96.7 (96.0, 97.4)	53.2 (45.0, 61.0)
7	2 OHIP/ NACRS claims in 1 year	Narrow	72.7 (63.8, 81.7)	97.8 (97.2, 98.4)	59.5 (50.6, 68.4)
		Broad	72.7 (63.8, 81.7)	97.8 (97.2, 98.4)	59.5 (50.6, 68.4)
8	2 OHIP/ NACRS claims in 2 years	Narrow	74.8 (66.0, 83.5)	97.8 (97.2, 98.4)	60.2 (51.4, 68.9)
		Broad	74.8 (66.0, 83.5)	97.8 (97.2, 98.4)	60.2 (51.4, 68.9)
9	2 OHIP/ NACRS claims in 3 years	Narrow	75.8 (67.2, 84.4)	97.8 (97.2, 98.4)	60.0 (51.3, 68.7)
		Broad	75.8 (67.2, 84.4)	97.8 (97.2, 98.4)	60.0 (51.3, 68.7)

Abbreviations: CI, confidence interval; CHF, congestive heart failure; CIHI, Canadian Institute for Health Information; NACRS, National Ambulatory Care Reporting System; OHIP, Ontario Health Insurance Plan; PPV, positive predictive value.

FIGURE 1
Results for nine algorithms to detect congestive heart failure (CHF) cases from administrative data using two CHF diagnostic code definitions plotted as on an ROC curve



Abbreviations: CHF, congestive heart failure; ROC, receiver operating characteristics.

consideration because their additional follow-up time is unnecessary). Based on the results of the summary measures, algorithm 3 appears superior. What's more, if this test were being used in a clinical setting it usually *would* be the best, primarily because its high sensitivity minimizes the number of cases that would be missed. However, for population-based reporting the false positive rate is a more important consideration. Because this algorithm will be used for an entire population, not just self-selected patients visiting a physician, and because the condition is relatively rare, even a small drop in specificity can translate into a large number of false positive cases. False positives, in turn, lead to research results that are biased to the null. As a result, algorithm 3 is ruled out by its relatively low specificity and LR+. The best algorithm for identifying cases of CHF is therefore algorithm 4.

In addition to testing the different algorithms, we also tested two different definitions of CHF. In most cases there

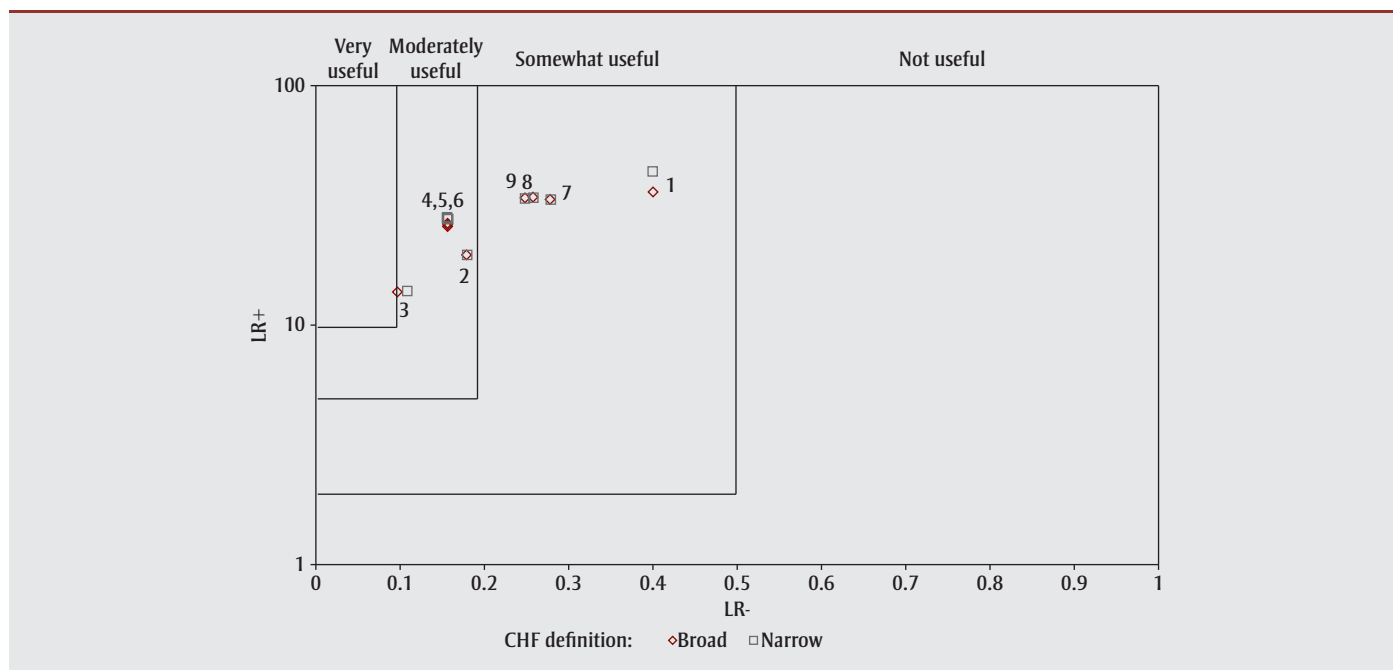
was little difference in the results. Where there was a difference, the narrow definition performed better, mainly because the broader definition tended to increase the number of false positives (Table 3).

Discussion

In this study, we tested nine different administrative data algorithms, which varied according to the number of records needed for a CHF diagnosis, the length of follow-up time allowed and the data sources used. We also tested two different sets of diagnostic codes used for identifying individuals with CHF in administrative data.

Previous studies that compared information in hospital charts with discharge data have found the coding of CHF diagnoses to be very accurate when tested against clinical criteria such as Framingham or Boston.^{4,5} However, we found that using hospital data alone to estimate CHF prevalence is insufficient and may fail to capture 40% of positive cases. This suggests that a sub-

FIGURE 2
Likelihood ratios (LR) for nine algorithms to detect congestive heart failure (CHF) cases from administrative data using two CHF diagnostic code definitions



Abbreviation: LR, likelihood ratio.

stantial proportion of Ontarians with CHF are being diagnosed and managed outside of hospital.

Another general finding is that basing a CHF diagnosis on only a single record with a CHF diagnostic code can successfully

identify individuals with the condition provided both hospitalization and ambulatory care data are used. However, this

TABLE 3
Comparison of congestive heart failure status from manually abstracted primary care physician data and administrative data using nine different algorithms (counts)

Algorithm	Description	CHF definition	True positive, N	False positive, N	True negative, N	False negative, N
1	1 CIHI record	Narrow	60	31	2208	39
		Broad	60	38	2201	39
2	1 OHIP claim	Narrow	82	95	2144	17
		Broad	82	95	2144	17
3	1 CIHI or OHIP claim	Narrow	89	145	2094	10
		Broad	90	149	2090	9
4	1 CIHI or 1 OHIP + 2nd claim (any source) in 1 year	Narrow	84	67	2172	15
		Broad	84	72	2167	15
5	1 CIHI or 1 OHIP + 2nd claim (any source) in 2 years	Narrow	84	68	2171	15
		Broad	84	73	2166	15
6	1 CIHI or 1 OHIP + 2nd claim (any source) in 3 years	Narrow	84	69	2170	15
		Broad	84	74	2165	15
7	2 OHIP/ NACRS claims in 1 year	Narrow	72	49	2190	27
		Broad	72	49	2190	27
8	2 OHIP/ NACRS claims in 2 years	Narrow	74	49	2190	25
		Broad	74	49	2190	25
9	2 OHIP/ NACRS claims in 3 years	Narrow	75	50	2189	24
		Broad	75	50	2189	24

Abbreviations: CHF, congestive heart failure; CIHI, Canadian Institute for Health Information; NACRS, National Ambulatory Care Reporting System; OHIP, Ontario Health Insurance Plan.

results in an unacceptably high number of false positives. It is possible that physicians may code CHF when ruling out CHF and a second diagnostic code would then be necessary to confirm that this is a true case of CHF.

We found the best algorithm for identifying cases of CHF to be one hospitalization record alone or one ambulatory care record if it is followed by a second record from any source within one year; this successfully identifies approximately 85% of patients with CHF, while keeping the false positive rate to the relatively low level of about 3%. This finding is similar to those in studies of other chronic conditions, such as diabetes,⁷ hypertension⁶ and ischemic heart disease,⁸ and in previous validation studies of the use of administrative data to identify CHF cases.^{10,11}

The fact that length of follow-up time to the second CHF record made little or no difference was somewhat surprising and differs from the results of similar studies for hypertension and diabetes.^{6,7} The explanation for this may lie with the fact that most people with true CHF are on medication and likely visiting a physician every few months. As a result, one year of follow-up is sufficient for our algorithm to detect CHF cases even if patients are not hospitalized. Hypertension and diabetes, on the other hand, may be initially managed without medication, which may lead to individuals going longer between visits to their physician, which in turn means that a longer follow-up time is required to pick up the second record.

The PPV for all algorithms seem somewhat low, ranging from 37.7 to 65.9. This is due to the fact that PPV is strongly related to prevalence: the lower the prevalence, the lower the PPV. CHF was relatively rare in our population, with a prevalence of only 4.3%; hence the low PPVs.

CHF is strongly age-related: extremely rare among those aged less than 40 years and rare among those aged between 40 and 65 years, it is increasingly prevalent after age 65 years. This prompts us to recommend that the algorithm not be used for populations where CHF is known to be very rare (i.e. among those aged less than 40 years).

The algorithm can be expected to perform very well among the elderly, particularly those aged over 75 years.

With respect to the ICD-9 and ICD-10 codes used to define CHF, we found no evidence that using an expanded list of diagnostic codes performed any better than the narrow definition of ICD-9 428 and ICD-10 I500, I501 and I509.

Limitations

The limitations to this study are mostly related to the fact that much of the data being used were collected for administrative purposes, not for research. A major limitation of the OHIP data is that only one diagnosis code is listed per billing. If CHF is not the patient's main reason for visiting a physician, it may not be recorded. Offsetting this limitation is the fact that Ontarians visit their physicians quite frequently, especially if they are aged over 65 years, thus providing a physician a number of chances to record CHF on a billing during a year. This is evident in the fact that length of follow-up made no difference in the ability of our algorithms to detect CHF cases.

While we recognize that we were unable to apply New York Heart Association (NYHA) diagnostic criteria for CHF to determine the presence of heart failure, our study used "real world" data based on physicians diagnosing and managing patients according to their own diagnostic acumen. While this may or may not fit into formal structured criteria, we would argue that if a physician is treating a patient as having CHF, it is reasonable to consider the patient as having the condition for incidence and prevalence reporting in a large population.

Conclusion

This study has shown that cases of CHF can be identified with a high degree of accuracy from administrative data, provided both ambulatory and hospitalization records are used. An algorithm of one hospitalization record, or one ambulatory record followed by a second record from either source within one year, with a diagnostic code definition of ICD-9 428 and ICD-10 I500, I501, I509, will identify

CHF patients with a sensitivity of 84.8% and a specificity of 97.0%.

Acknowledgements

This study was supported by the Institute for Clinical Evaluative Sciences (ICES), which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care. The opinions, results and conclusions reported in this paper are those of the authors and are independent from the funding sources. No endorsement by ICES or the Ontario Ministry of Health and Long-Term Care is intended or should be inferred.

Dr. Karen Tu is supported by a Fellowship in Primary Care Research by the Canadian Institutes of Health Research (CIHR).

This study was funded by a grant from the Public Health Agency of Canada and by a Canadian Institutes of Health Research team grant in Cardiovascular Outcomes Research to the Canadian Cardiovascular Outcomes Research Team (CCORT).

Conflict of interest: none.

References

1. Lee DS, Johansen H, Gong Y, et al. Regional outcomes of heart failure in Canada. *Can J Cardiol.* 2004;20(6):599-607.
2. Hall RE, Tu JV. Hospitalization rates and length of stay for cardiovascular conditions in Canada, 1994 to 1999. *Can J Cardiol.* 2003;19(10):1123-31.
3. Saczynski JS, Andrade SE, Harrold LR, et al. A systematic review of validated methods for identifying heart failure using administrative data. *Pharmacoepidemiol Drug Saf.* 2012;21(S1):129-40.
4. Lee DS, Donovan L, Austin PC, et al. Comparison of coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research. *Med Care.* 2005;43(2):182-8.
5. Teng TH, Finn J, Hung J, et al. A validation study: how effective is the Hospital Morbidity Data as a surveillance tool for heart failure in Western Australia? *Aust N Z J Public Health.* 2008;32(5):405-7.

6. Tu K, Campbell NR, Chen ZL, et al. Accuracy of administrative databases in identifying patients with hypertension. *Open Medicine*. 2007;1(1):E18-26.
7. Hux JE, Ivis F, Flintoft V, et al. Diabetes in Ontario: determination of prevalence and incidence using a validated administrative data algorithm. *Diabetes Care*. 2002;25(3):512-6.
8. Tu K, Mitiku T, Lee DS, et al. Validation of physician billing and hospitalization data to identify patients with ischemic heart disease using data from the Electronic Medical Record Administrative data Linked Database (EMRALD). *Can J Cardiol*. 2010;26(7):e225-8.
9. Gershon AS, Wang C, Guan J, et al. Identifying individuals with physician diagnosed COPD in health administrative databases. *COPD*. 2009;6(5):388-94.
10. Alqaisi F, Williams LK, Peterson EL, et al. Comparing methods for identifying patients with heart failure using electronic data sources. *BMC Health Serv Res*. 2009;9:237.
11. McCullough PA, Philbin EF, Spertus JA, et al. Confirmation of a heart failure epidemic: findings from the Resource Utilization Among Congestive Heart Failure (REACH) study. *J Am Coll Cardiol*. 2002;39(1):60-9.
12. Chan BT, Schultz SE. Supply and utilization of general practitioner and family physician services in Ontario. ICES Investigative Report. Toronto (ON): Institute for Clinical Evaluative Sciences; 2005.
13. Vermeulen MJ, Tu JV, Schull MJ. ICD-10 adaptations of the Ontario acute myocardial infarction mortality prediction rules performed as well as the original versions. *J Clin Epidemiol*. 2007;60(9):971-4.
14. Li B, Evans D, Faris P, et al. Risk adjustment performance of Charlson and Elixhauser comorbidities in ICD-9 and ICD-10 administrative databases. *BMC Health Serv Res*. 2008;8:12.
15. Chen G, Faris P, Hemmelgarn B, et al. Measuring agreement of administrative data with chart data using prevalence unadjusted and adjusted kappa. *BMC Med Res Methodol*. 2009; 9:5.
16. Ontario Health Technology Advisory Committee. Multi-detector computed tomography angiography for coronary artery disease. Toronto (ON): Ministry of Health and Long-Term Care; 2005.
17. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA*. 1994;271(9):703-7.
18. Statistics Canada. Annual Demographic Estimates: Canada, Provinces and Territories. Ottawa (ON); Statistics Canada;2009 [Statistics Canada, Catalogue No.: 91-215-X].
19. Bertakis KD, Azari R, Helms LJ, et al. Gender differences in the utilization of health care services. *J Fam Pract*. 2000;49(2):147-52.