

# Validation of Administrative Databases in Identifying Patients with Hypertension

Karen Tu MD, MSc<sup>a,b</sup>, Norman RC Campbell MD<sup>c</sup>, Zhong-liang Chen MD, BSc, MSc<sup>a</sup>, Karen J Cauch-Dudek BA<sup>a</sup>, Finlay A McAlister MD, MSc<sup>d</sup>

<sup>a</sup>Institute For Clinical Evaluative Sciences (ICES), Toronto, Canada, <sup>b</sup>Department of Family and Community Medicine, University of Toronto, Toronto, Canada

<sup>c</sup>Departments of Medicine and Pharmacology and Therapeutics, University of Calgary, Calgary, Canada, <sup>d</sup>Division of General Internal Medicine, University of Alberta, Edmonton, Canada,

This research was supported by a Grant-in-Aid from the Heart and Stroke Foundation of Ontario Grant #NA 5459. KT is supported by a Canadian Institutes of Health Research (CIHR) Short-Term Clinician Investigator Award. FM is supported by an Alberta Heritage Foundation for Medical Research Population Health Scholar Award, a CIHR New Investigator Award, and the University of Alberta/Merck Frosst/Aventis Chair in Patient Health Management.

NC has been paid fees for speaking and for consulting by most pharmaceutical companies in Canada that produce trade medications to lower blood pressure. In addition, NC has received research funds from Pfizer Canada, Sanofi-Aventis, Merck Frosst, Servier and Bristol-Myers-Squibb for research on the epidemiology of hypertension. NC also chairs the steering committee and executive committee of the Canadian Hypertension Education Program (CHEP) as a volunteer. CHEP has a mandate to improve the treatment and control of hypertension in Canada. FM has received operating grant funding from Pfizer Canada for an ongoing investigator-initiated trial co-funded by the Heart and Stroke Foundation of Canada and Pfizer Canada. ZC, KC and KT have no potential conflicts of interest.

## Corresponding Author

Dr Karen Tu  
c/o ICES  
G108, 2075 Bayview Avenue  
Toronto, Ontario  
M4N 3M5  
Canada

Phone (416) 480-4055 x3871  
Fax (416) 480-6048  
Email karen.tu@ices.on.ca

**Word Count: 3419**

## **Abstract**

**Objectives** To determine the accuracy of using administrative databases to identify people with hypertension in Ontario, Canada's largest province.

**Methods** Data on hypertension status in adults was abstracted from a random sample of primary care physician offices throughout the province and used as the reference standard to judge the accuracy of using different types of administrative data to identify cases of hypertension.

**Results** A case definition algorithm employing two outpatient physician billing claims over a three year period had a sensitivity of 73.0% (95% confidence interval [CI] 69.2-76.8), a specificity of 94.7% (CI 93.4-96.0), a positive predictive value of 86.7% (CI 89.6-89.9), and a negative predictive value of 88.2% (CI 89.6-90.0) for detecting hypertensive adults compared to physician-assigned diagnoses. While the same case definition compared to self-reported survey data had a sensitivity of 64.2% (CI 63.0-65.5), a specificity of 93.8% (CI 93.4-94.2), a positive predictive value of 77.2% (76.0-78.4), and a negative predictive value of 88.9% (CI 88.5-89.4). Applying this algorithm to the entire province of Ontario, the age and sex standardized prevalence of hypertension in adults older than 35 years increased from 20% in 1994 to 31% in 2004.

**Conclusions** It is possible to use administrative data to accurately identify patients with hypertension. The use of administrative data is a cheaper and more feasible method for surveillance of hypertension occurrence and outcomes over time in a large population than serial cross-sectional or cohort studies.

**Key Words** hypertension, administrative databases, surveillance

Traditionally, surveillance for occurrence and outcomes of chronic conditions largely dealt with in the outpatient setting, such as hypertension or diabetes mellitus, have been based on self-report surveys, chart audits (generally of a relatively small number of physicians from selected locales), or physical measure surveys. However, none of these approaches can be used for an ongoing national chronic disease surveillance program as they are time consuming, costly, and generally are conducted at single points of time or in restricted locales with highly select samples. Although self report surveys are feasible to repeat in a longitudinal fashion in a broader array of individuals, attrition of subjects can be a problem and the accuracy of self-reported diagnoses for conditions such as hypertension have been called into question.[1, 2] The most recent estimates of hypertension prevalence based on physical measures in Canada are over 10 years old.[3]

However, Canada has a universal healthcare system and single provincial government payer – as a result, administrative claims data (with diagnoses) are routinely collected for medical services, including hospitalizations, most physician visits, and (in some provinces) prescription medications. Linking these various administrative claims databases can provide an inexpensive population-based, longitudinal record of patient care independent of treatment site and with virtually 100% capture without problems due to recall bias, non-response, or losses to follow-up (thus addressing the key concerns with the current surveillance approaches discussed above). Indeed, the use of these administrative claims databases for identifying certain diseases, their treatments, and outcomes has been well studied for conditions largely dealt with in the hospital setting (such as myocardial infarction, heart failure, and stroke).[4-7] While the use of outpatient physician billing records to explore the occurrence of chronic ambulatory conditions is less well studied, it has

recently been established that administrative claims data can accurately identify patients with diabetes (with a sensitivity of 86% and a positive predictive value of 80%) – as a result, in Canada there is now a national diabetes surveillance program based entirely on administrative claims data (ie. an amalgam of physician claims data and hospitalization diagnostic code data).[8] Whether such a system could work for hypertension surveillance on a national basis (both in Canada and in other countries that may routinely collect similar type data) depends on the accuracy of administrative claims data in identifying individuals with hypertension.

Previous studies have looked at agreement of hypertension diagnoses between physician billing claims and self-report survey data,[9] clinical measures data,[10] chart audits in elderly patients,[11] medication dispensations,[12] and American Health insurance medical and pharmacy claims.[13] However, these studies have been limited by restriction to non-random study samples or elderly patients only, or by examination of claims data for short periods of time only (such as 12 months), or by examination of single types of claims data in isolation (for example, outpatient physician billing claims alone or inpatient hospitalization claims alone rather than both together) or occurred in other jurisdictions outside of Canada. As a result, the accuracy of using administrative claims data in adult patients of all ages and the usefulness of amalgamating hospitalization data with outpatient physician claims data to identify cases of hypertension remains unclear. In addition, the implications of different case definition algorithms on the accuracy of hypertension diagnosis and estimates of population prevalence and incidence for hypertension are unknown.

Thus, we designed this study to examine the accuracy (against a variety of reference standards) and impact of different administrative database case definition

algorithms for hypertension surveillance in Ontario (Canada's largest province with a total population of over 12 million individuals in 2004).

## METHODS

### ***Data Sources for Validation of Administrative Data Algorithm:***

#### *1. Outpatient Primary Care Chart Data:*

Between December 2004 and August 2005 we audited patient charts from 76 fee-for-service family physician volunteers (from throughout Ontario and all of whom had been in their current practice site for at least three years) and used charted diagnoses as the reference standard for comparison of hypertension status with administrative data. We abstracted data from a random sample of 1676 charts of adult patients that fit our inclusion criteria of: age 35 years or older as of three years prior to the abstraction date, they were regular patients of the participating physician (as defined as at least two of: seen the most, did a complete physical, or registered under the participating physician), the patient was still in the practice, the date of the first visit to the participating physician was at least three years prior, there were at least 2 visits in the prior three years and their healthcard number was available. We collected healthcard numbers and data from progress notes, lab results and consult notes three years prior to the abstraction date and also from cumulative patient profiles where available and classified individuals as having a diagnosis of hypertension if they had a physician-assigned diagnosis recorded in their chart, they had a prescription for an antihypertensive in the context of an elevated blood pressure reading, or if their recorded blood pressures met the criteria for diagnosis laid out in the Canadian Hypertension Education Program guidelines (which were stable throughout the years our chart audit data were drawn from).[14] Our sample size was pre-chosen based on an estimated hypertension prevalence of 20% in adults 35 years or older and 40% in the 65 and over age group[3], an assumed sensitivity of 78%[10]

with standard error calculation of  $\pm 0.5$ , and a 25% inflation factor in case our prevalence was lower than estimated. Ten percent of the charts were abstracted by two abstractors and overall agreement for the presence or absence of hypertension was very good at 92.8% agreement (kappa 0.84).

## 2. *Administrative Claims Data:*

Using the primary care chart abstracted data as the reference standard, the healthcard numbers that were collected were encrypted and converted into unique identifiers and linked to the Ontario Health Insurance Plan (OHIP) physician billing claims database and the Canadian Institute for Health Information (CIHI) hospitalization discharge abstracts database for each patient in the chart audit study. Linkage was conducted for the 3 years preceding the chart abstraction date and claims for hypertension specific codes (ICD-9-CM codes 401.x, 402.x, 403.x, 404.x, or 405.x and/or ICD-10 codes I10.x, I11.x, I12.x, I13.x, or I15.x in CIHI fiscal years 2002-2004) in any of the 16 diagnostic fields included in CIHI data or the 1 diagnostic field included in OHIP data for any billing physician (not just the physicians whose charts we abstracted data from) were flagged and used to explore the percent agreement, kappa, sensitivity, specificity, and positive/negative predictive values between primary care chart audit and administrative claims data for 15 different case definition algorithms for hypertension (see Table 1 for description of case definition algorithms). The chart audit diagnosis was used as the reference standard and 95% confidence intervals were calculated using a binomial probability distribution. We defined sensitivity as the proportion of individuals with hypertension documented in their physician chart who were identified as having hypertension using administrative data and we defined specificity as the proportion of individuals without hypertension documented in their physician chart who were identified as not having hypertension

using the administrative data. We defined positive predictive value as the proportion of individuals identified as having hypertension in the administrative data whose diagnosis was confirmed by chart audit and negative predictive value as the proportion of individuals identified as not having hypertension in the administrative data whose lack of a hypertension diagnosis was confirmed by chart audit.

*Sensitivity Analyses:*

(1) We also tested the same 15 hypertension case definition algorithms in the administrative claims data against a self-reported survey reference standard - the Canadian Community Health Survey 2001.[15] The data for this self-report survey was collected between September 2000 and November 2001 and included 39, 278 respondents from Ontario aged 12 and over of whom 84% agreed to linkage with administrative data for research purposes. Patients were defined as being hypertensive if they answered yes to either of two questions- “do you have high blood pressure?” or “in the past month did you take medicine for blood pressure?”

(2) To explore the performance of these case definition algorithms in different patient populations, we planned 2 additional sensitivity analyses. In the first, we tested the performance characteristics of each hypertension case definition algorithm in the elderly subgroup from our primary care chart audit (ie. those patients aged over 65 as of March 31, 2001, three years prior to the most recent hospital discharge abstract data available). In the second, we evaluated the performance characteristics of each hypertension case definition algorithm in a cohort of 20,830 cardiac patients who were discharged alive after a hospitalization for myocardial infarction or heart failure from 1999/00 to 2000/01 in 103 acute care facilities in Ontario (the Enhanced Feedback for Effective Cardiac Treatment [EFFECT] study cohort).[16] Full details about the EFFECT study, including patient and hospital



demographics as well as variable definitions and data collection procedures are available at <http://www.ccort.ca/EFFECT.asp>.

***Determination of Rates of Hypertension:***

We examined the impact of different administrative database case definition algorithms on estimates of the prevalence of hypertension in the adult population of Ontario, Canada in fiscal year 2004. In addition, we examined annual population adjusted, age and sex standardized changes, according to 2004 Statistics Canada census population records for Ontario, in the prevalence and incidence of hypertension between 1994 and 2004 in Ontario using all 15 case definition algorithms (but only report the results for the two physician claims in 3 years algorithm in this manuscript - details of other analyses available from Dr. Tu directly). Prevalent cases were carried forward for each year, patients that died or moved out of the province were excluded, and only individuals with no previous claims for hypertension were counted as incident cases for the relevant fiscal year. Although we have OHIP administrative data from 1991 onward, we did not examine prevalence and incidence rates until the 1994 fiscal year in order to have three prior years of physician billing data to use as a “wash-out period” in order to minimize the possibility that we would overestimate incidence rates in the initial years of our study. *A priori* we recognized that using a case definition algorithm requiring 2 claims in 3 years would slightly underestimate incident cases in 2003 and 2004 given the absence of hospitalization data after fiscal 2004 and physician billing data after fiscal 2005 respectively. We also examined age standardized prevalence and incidence of hypertension by gender.

## RESULTS

Our chart audit study sample consisted of 1676 adult patients (average age 55.6 years) with 32% (547 patients) having a diagnosis of hypertension. As expected, the prevalence of hypertension in our sub-sample of 1038 patients older than 65 years (average age 74.7 years) was higher at 63% (653 patients). The prevalence of hypertension in the 20,830 EFFECT study patients (patients hospitalized for cardiovascular disease 1999/00 to 2001/01, average age 71.6 years) was 46% whereas the prevalence in the 22,087 individuals age 35 and over in the 2001 Canadian Community Health Survey (average age 55.6 years) was 25%.

The overall agreement between the primary care chart diagnosis and administrative claims data was greater than 80% for all but one of the hypertension case definition algorithms (Table 1). Clearly, using a diagnosis of hypertension in a hospital discharge summary was far less sensitive than any other administrative data case definition algorithm approach (but was highly specific) – Table 1. As one would expect, those case definition algorithms which required more than a single claim for hypertension in the administrative data before classifying an individual as being hypertensive had higher specificity and positive predictive values than definitions based on a single billing or hospitalization claim (Table 1). Also as expected, those case definition algorithms which were based on a longer observation period had a greater sensitivity for the detection of hypertension than those algorithms based on the administrative data from a single year. Receiver Operating Characteristic (ROC) Curve analyses demonstrated that the case definitions using 2 physician outpatient billing claims for hypertension in administrative data with the greatest area under the curve were (i) 2 physician outpatient billing claims in 3 years or (ii) 2 physician outpatient billing claims or 1 hospital discharge in back to 1988. Given the time lag in

obtaining hospital discharge data in many administrative datasets (in Ontario the lag can be as long as 2 fiscal years for hospital discharge data while the physician outpatient billing data lag is approximately 3 months), we chose the “2 physician billing claims in 3 years” as the case definition algorithm for our future surveillance work on hypertension in Ontario.

#### *Sensitivity Analyses:*

Comparison to self-reported diagnoses of hypertension in the 2001 Canadian Community Health Survey (Table 2) confirmed that the “2 physician billing claims in 3 years” case definition algorithm for hypertension was also highly accurate in that dataset, with sensitivities and specificities similar to those seen when the primary care chart diagnosis was used as the reference standard. This case definition algorithm also performed well in older patients from our primary care chart audit (an important subgroup to examine since many pharmacoepidemiologic studies are restricted to the over 65 age group) – Table 3. However, none of the hypertension case definition algorithms performed well in the EFFECT cohort of cardiac patients discharged after hospitalization for heart failure or myocardial infarction (Table 3). The poor sensitivity of administrative claims data for detecting hypertension in these patients with substantial co-morbidities is not surprising since the outpatient billing claims only require one diagnosis and it is thus likely that in a patient recently discharged after an acute cardiac condition such as myocardial infarct or heart failure that that diagnosis would predominate and be the one specified in follow-up appointments, even if the patient had hypertension

#### *Hypertension occurrence rates over time:*

Using our “2 physician outpatient claims in 3 years” case definition, the age and sex standardized prevalence of hypertension in Ontario rose steadily from 20% of the

population aged 35 or older in 1994 to 31% in 2004; however, the age and sex standardized incidence remained relatively constant at approximately two percent of the population per year (Figure 1). Prevalence of hypertension was higher in females than males in Ontario during the years of our study, although incidence rates were similar between genders (Figure 2) – this is not unexpected given the longer survival of women, the higher prevalence of hypertension in females with age relative to men and the greater awareness of hypertension in women compared to men.[17]

## DISCUSSION

We examined multiple different algorithms for defining hypertension in administrative data and present the accuracy of each algorithm validated against both physician-assigned diagnosis (in our audit of adults from primary care practices) and against self-reported diagnosis (in a survey of community-dwelling adults). We have demonstrated that the use of administrative data to define hypertension and conduct ongoing surveillance of prevalence and incidence is feasible and reasonably accurate in adult patients, including those older than 65 years old that likely have more multiple co-morbidities. However, our analyses demonstrating substantially lower accuracy of administrative claims data for hypertension when compared with physician-assigned diagnoses in the cardiac patients of the EFFECT cohort suggests that the use of administrative data for hypertension surveillance is best suited to patients with uncomplicated hypertension and not suitable for patients with substantial cardiac co-morbidities.

The accuracy of the hypertension case definition algorithm we suggest (“2 [or more] physician outpatient billing claims in 3 years”) compares very favorably with the diabetes case definition algorithm used currently in Canada for a national diabetes surveillance program (“2 physician outpatient billing claims or 1 hospital discharge in 2 years”).[8] Indeed, the positive predictive value of our hypertension algorithm (87%) exceeds the 80% positive predictive value for the diabetes algorithm which has become the “gold standard” for diabetes health outcomes research employing administrative databases. However, it should be recognized that there is always a trade-off between sensitivity and specificity in choosing a case definition algorithm for any condition in an administrative claims database and the optimal algorithm for defining hypertension depends upon the purpose of a study. Thus, when defining the

burden of illness with hypertension for the purposes of resource allocation planning, one would conceivably wish to identify as many cases as possible and thereby choose an algorithm with the highest sensitivity. On the other hand, in examining practice patterns and outcomes from hypertension, one would wish to choose an algorithm with the highest specificity to ensure that as few non-hypertensives as possible are included in the study sample. Thus, the data we provide in Tables 1 and 2 will have wider application for hypertension outcomes researchers than just establishing the best case definition algorithm for a national surveillance program.

Our findings confirm and extend the data from previous studies examining the detection of hypertension using administrative databases. For example, similar to a previous study,[9] we found that the use of hospitalization data did not greatly enhance the accuracy of administrative data in identifying people with hypertension. Further, the sensitivity and specificity of the “one physician billing claim” case definition algorithm was similar in our datasets to the results reported in two other Canadian provinces where this case definition algorithm was tested - Manitoba[10] (one physician billing claim in two years) and Quebec[11] (one physician billing claim in one year). However, our study extends those studies by examining 14 other hypertension case definition algorithms in administrative data and by testing the algorithms in a variety of patient groups and for both physician-assigned diagnosis and self-reported diagnosis.

Our finding of a slightly higher discordance between hypertension case definition using administrative data and self-reported survey diagnoses is not surprising since approximately 5% of persons reporting drug treatment for hypertension do not report a diagnosis of hypertension in Canadian surveys[1] (presumably because they erroneously think that their condition is cured if they are

taking medication). Thus, akin to investigators in diabetes health outcomes research, we believe that self-reported diagnosis is inadequate as a gold standard for hypertension and thus place primary emphasis on our administrative data validation work with the physician-assigned diagnosis in the primary care chart audit.

There are some limitations to our study. Our identification of patients with hypertension was limited to only those patients that visited physicians. However, in Ontario health care is paid for by the government and is free to patients- thus, less than 6% of Ontarians report not having a family physician and less than 25% do not visit a primary care physician at least once a year. Thus, within a two or three year time period it is likely that almost all Ontarians have at least one visit to a physician. [18] Indeed, in the 2001 Canadian Community Health Survey, 73% of respondents reported having their blood pressure measured within the past year and 85% reported having a blood pressure measured in the 2 years before the survey.[15] Nonetheless, we do acknowledge that our prevalence estimates may underestimate the true prevalence of hypertension as it is unknown exactly how many cases of hypertension go undetected in Canada. Although the majority of the physicians in Ontario bill under the OHIP billing plan, approximately 2% of physicians are paid salaries under alternate funding plans such that they do not bill fee-for-service and therefore their activity does not show up on the OHIP physician billing claims data.[18] This in turn may also lead to a slight underestimate of the occurrence of hypertension within the population of Ontario. Finally, we acknowledge that the generalizability of our hypertension coding algorithm in other jurisdictions with different administrative data is unknown, but studies are ongoing to validate our findings in other jurisdictions.

While the ideal study in which to examine hypertension occurrence and outcomes would be a prospective cohort study with assessment of actual blood

pressure measurements (and other physiologic parameters) at repeated intervals in tens of thousands of patients, such a study would be prohibitively expensive and would never be a practical option for ongoing surveillance of hypertension occurrence and outcomes. Indeed, the last national cross sectional study that assessed blood pressure in Canada (the Canadian Heart Health Survey) was conducted in the early 1990s and has not yet been repeated due to financial restraints. Using administrative data to study hypertension occurrence and outcomes holds several potential attractions, not least of which include their readiness to be analyzed, their wide geographic coverage, and their relatively complete capture of episodes of patient contact with the health system. Our study has established that administrative data can be used to accurately examine the occurrence of hypertension.

**Acknowledgements** Thank you to Jack V. Tu for supplying us with EFFECT study data and Janet E. Hux for reviewing the manuscript.



## Reference List

1. Gentleman JF, Tomiak M. The Consistency of Various High Blood Pressure Indicators Based on Questionnaire and Physical Measures Data from the Canada Health Survey. *Health Reports*. vol. 4, no 3, 293-311. 1992. Statistics Canada.
2. Wagner EH, James SA, Beresford SA, Strogatz DS, Grimson RC, Kleinbaum DG, et al. The Edgecombe County High Blood Pressure Control Program: I. Correlates of Uncontrolled HYPertension at Baseline. *Am J Public Health* 1984; 74:237-242.
3. Joffres MR, Hamet P, Rabkin SW, Gelskey D, Hogan K, Fodor G. Prevalence, control and awareness of high blood pressure among Canadian adults. Canadian Heart Health Surveys Research Group. *CMAJ* 1992; 146(11):1997-2005.
4. Tu J, Austin P, Naylor C, Iron K, Zhang H. Acute Myocardial Outcomes in Ontario. In: Naylor C, Slaughter P (editors): *Cardiovascular Health and Services in Ontario: An ICES Atlas*. Toronto: Institute for Clinical Evaluative Sciences; 1999, pp. 83-110.
5. Jong P, Gong Y, Liu PP, Austin PC, Lee DS, Tu JV. Care and Outcomes of Patients Newly Hospitalized for Heart Failure in the Community Treated by Cardiologists Compared With Other Specialists. *Circulation* 2003; 108:184-191.
6. Hill MD, Yiannakoulis N, Jeerakathil T, Tu JV, Svenson LW, Schopflocher DP. The high risk of stroke immediately after transient ischemic attack: a population-based study. *Neurology* 2004; 62(11):2015-2020.
7. Austin PC, Daly PA, Tu JV. A multicenter study of the coding accuracy of hospital discharge administrative data for patients admitted to cardiac care units in Ontario. *Am Heart J* 2002; 144(2):290-296.
8. Hux JE, Ivis F, Flintoft V, Bica A. Diabetes in Ontario: determination of prevalence and incidence using a validated administrative data algorithm. *Diabetes Care* 2002; 25(3):512-516.
9. Robinson JR, Young TK, Roos LL, Gelskey DE. Estimating the burden of disease. Comparing administrative data and self-reports. *Med Care* 1997; 35(9):932-947.
10. Muhajarine N, Mustard C, Roos LL, Young TK, Gelskey DE. Comparison of survey and physician claims data for detecting hypertension. *J Clin Epidemiol* 1997; 50(6):711-718.
11. Wilchesky M, Tamblyn RM, Huang A. Validation of diagnostic codes within medical services claims. *J Clin Epidemiol* 2004; 57(2):131-141.
12. Quam L, Ellis LB, Venus P, Clouse J, Taylor CG, Leatherman S. Using claims data for epidemiologic research. The concordance of claims-based criteria with the medical record and patient survey for identifying a hypertensive population. *Med Care* 1993; 31(6):498-507.

13. Bullano MF, Kamat S, Willey VJ, Barlas S, Watson DF, Brenneman SK. Agreement Between Administrative Claims and the Medical Record in Identifying Patients With a Diagnosis of Hypertension. *Med Care* 2006; 44(5):486-490.
14. Hemmelgarn BR, McAllister FA, Myers MG, McKay DW, Bolli P, Abbott C, et al. The 2005 Canadian Hypertension Education Program recommendations for the management of hypertension: Part 1 - Blood pressure measurement, diagnosis and assessment of risk. *Can J Cardiol* 2005; 21(8):645-656.
15. Canadian Community Health Survey. 2001. Ottawa, Statistics Canada, Health Statistics Division.
16. Tu JV, Donovan LR. Quality of Cardiac Care in Ontario. EFFECT (Enhanced Feedback for Feedback for Effective Cardiac Treatment) Phase I Report I. 2004. The Institute for Clinical Evaluative Sciences (ICES).
17. Hajjar I, Kotchen TA. Trends in Prevalence, Awareness, Treatment and Control of Hypertension in the United States, 1988-2000. *JAMA* 2003; 290(2):199-206.
18. Chan BTB, Schultz SE. Supply and Utilization of General Practitioner and Family Physician Services in Ontario. ICES Investigative Report. 43. 2005. Toronto, Institute for Clinical Evaluative Sciences.

**Table 1: Validation of hypertension case definition algorithms against primary care chart data for 1676 patients older than 35 years (32% with chart diagnosis of hypertension)**

Case Definition Algorithm	Percent Agreement	Kappa	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Hypertension prevalence in Ontario adults (2004 census) using each definition
<b>Using Outpatient Physician Billing (OHIP) Claims Only:</b>							
1 claim in one year	84.3(82.6-86.1)	0.62(0.58-0.66)	64.8(60.8-68.8)	93.5(92.1-94.9)	82.5(78.8-86.1)	84.9(82.9-86.9)	42.8%
1 claim in two years	86.4(84.8-88.0)	0.69(0.65-0.72)	78.2(74.7-81.7)	90.3(88.5-92.0)	79.1(75.6-82.6)	89.8(88.0-91.5)	42.8%
1 claim in three years	86.9(85.3-88.5)	0.71(0.67-0.74)	83.6(80.5-86.7)	88.4(86.6-90.3)	77.3(73.9-80.7)	92.0(90.4-93.6)	42.8%
2 claims in one year	80.7(78.8-82.6)	0.49(0.44-0.53)	43.9(39.7-48.1)	98.1(97.3-98.9)	91.5(88.1-94.9)	78.8(76.6-80.9)	30.1%
2 claims in two years	86.0(84.4-87.7)	0.65(0.61-0.69)	64.1(60.0-68.1)	96.4(95.3-97.5)	89.4(86.3-92.4)	85.1(83.1-87.0)	31.4%
2 claims in three years	87.8(86.2-89.3)	0.71(0.67-0.74)	73.0(69.2-76.8)	94.7(93.4-96.0)	86.7(83.6-89.9)	88.2(86.3-90.0)	32.1%
<b>Using Outpatient Physician Billing (OHIP) Claims and Inpatient Hospitalization (CIHI) Data:</b>							
1 claim or 1 discharge in one year	84.8(83.1-86.6)	0.63(0.59-0.67)	67.0(63.1-71.0)	93.2(91.8-94.7)	82.4(78.8-86.0)	85.7(83.8-87.7)	43.4%
1 claim or 1 discharge in two years	86.9(85.3-88.5)	0.70(0.66-0.74)	80.3(76.9-83.6)	90.0(88.2-91.7)	79.1(75.7-82.5)	90.6(88.9-92.3)	43.4%
1 claim or 1 discharge in three years	87.4(85.8-88.9)	0.72(0.68-0.75)	85.3(82.3-88.3)	88.3(86.5-90.2)	77.5(74.1-80.9)	92.7(91.2-94.3)	43.4%
2 claims or 1 discharge in one year	83.9(82.1-85.7)	0.59(0.55-0.63)	55.9(51.7-60.1)	97.1(96.1-98.1)	90.1(86.9-93.3)	82.4(80.3-84.4)	31.3%
2 claims or 1 discharge in two years	87.7(86.1-89.3)	0.70(0.67-0.74)	72.3(68.5-76.0)	95.0(93.7-96.3)	87.2(84.1-90.3)	87.9(86.1-89.7)	32.5%
2 claims or 1 discharge in three years	88.1(86.6-89.7)	0.72(0.69-0.76)	77.5(73.9-81.0)	93.2(91.7-94.6)	84.2(81.0-87.4)	89.8(88.0-91.5)	33.1%
2 claims in 2 years or 1 discharge back to 1988	87.7(86.1-89.3)	0.71(0.67-0.74)	73.0(69.2-76.8)	94.6(93.3-96.0)	86.5(83.4-89.7)	88.1(86.3-90.0)	32.6%
2 claims in 3 years or 1 discharge back to 1988	88.1(86.6-89.7)	0.72(0.69-0.76)	78.2(74.7-81.7)	92.8(91.3-94.3)	83.7(80.4-86.9)	90.0(88.3-91.7)	33.2%
1 discharge in three years	70.3(68.1-72.5)	0.10(0.07-0.13)	8.0(5.7-10.3)	99.6(99.3-100.0)	91.5(83.5-99.5)	69.7(67.4-71.9)	8.4%

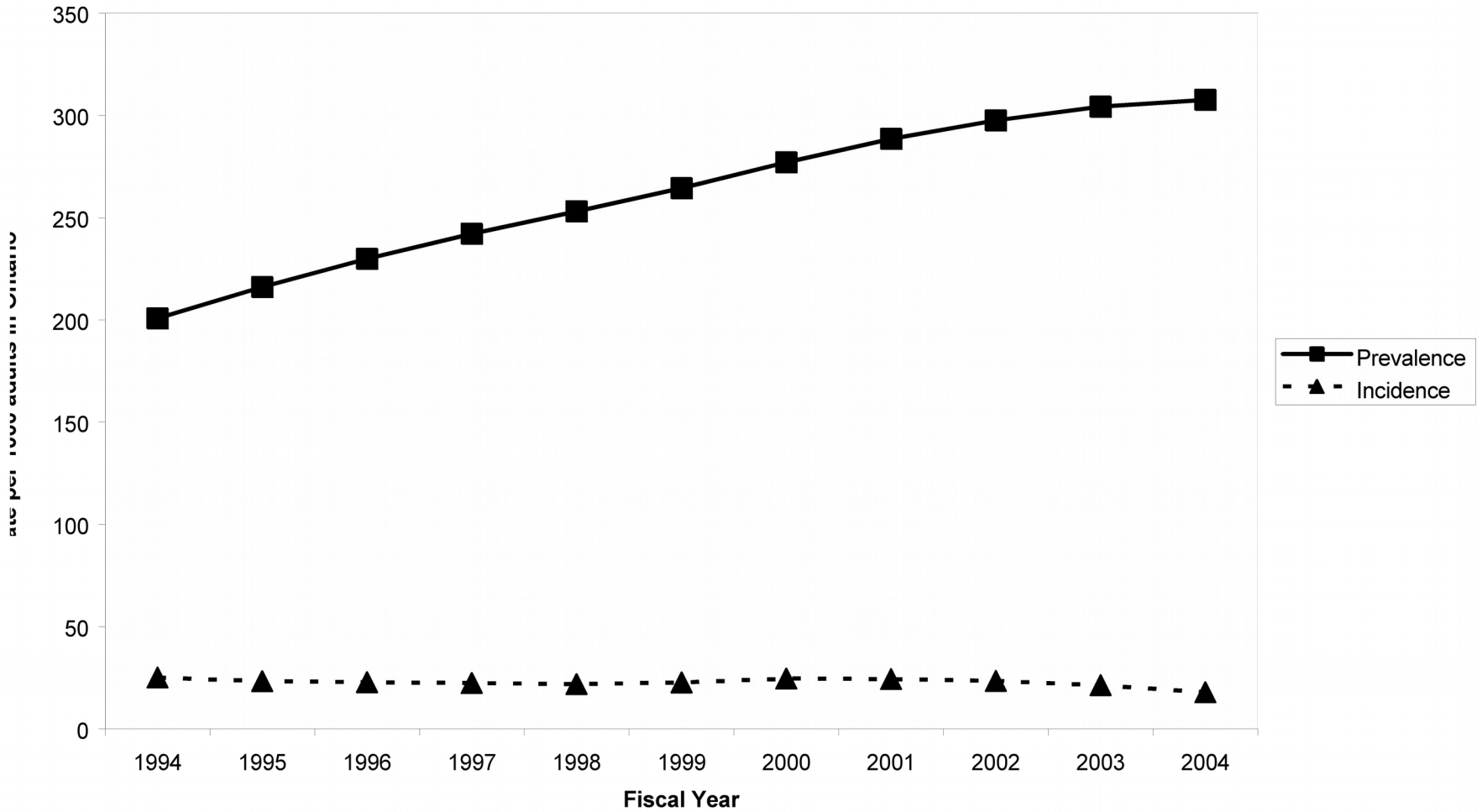
**Table 2: Validation of hypertension case definition algorithms against self-report survey data for 22 087 adult patients  
(23% with self-reported diagnosis of hypertension)**

<b>Case Definition Algorithm</b>	<b>Percent Agreement</b>	<b>Kappa</b>	<b>Sensitivity</b>	<b>Specificity</b>	<b>Positive Predictive Value</b>	<b>Negative Predictive Value</b>
<b>Using Outpatient Physician Billing (OHIP) Claims Only:</b>						
1 claim in one year	85.1(84.6-85.5)	0.57(0.55-0.58)	59.2(57.9-60.5)	93.5(93.1-93.9)	74.9(73.6-76.2)	87.5(87.0-88.0)
1 claim in two years	85.1(84.7-85.6)	0.60(0.59-0.61)	70.5(69.3-71.7)	89.9(89.4-90.4)	69.5(68.3-70.7)	90.3(89.9-90.8)
1 claim in three years	84.5(84.0-84.9)	0.61(0.59-0.61)	76.5(75.3-77.6)	87.1(86.6-87.6)	65.9(64.7-67.1)	91.9(91.5-92.3)
2 claims in one year	83.6(83.1-84.0)	0.46(0.45-0.48)	40.6(39.3-41.9)	97.6(97.3-97.8)	84.5(83.1-85.9)	83.4(82.9-83.9)
2 claims in two years	85.9(85.4-86.3)	0.58(0.56-0.59)	56.4(55.1-57.8)	95.5(95.1-95.8)	80.3(79.0-81.5)	87.0(86.5-87.5)
2 claims in three years	86.5(86.1-87.0)	0.62(0.60-0.62)	64.2(63.0-65.5)	93.8(93.4-94.2)	77.2(76.0-78.4)	88.9(88.5-89.4)
<b>Using Outpatient Physician Billing (OHIP) Claims and Inpatient Hospitalization (CIHI) Data:</b>						
1 claim or 1 discharge in one year	85.2(84.7-85.7)	0.57(0.56-0.57)	60.3(59.0-61.6)	93.3(92.9-93.7)	74.7(73.4-76.0)	87.8(87.3-88.3)
1 claim or 1 discharge in two years	85.2(84.8-85.7)	0.61(0.59-0.62)	71.7(70.5-72.9)	89.7(89.2-90.1)	69.4(68.2-70.6)	90.7(90.2-91.1)
1 claim or 1 discharge in three years	84.6(84.1-85.1)	0.61(0.60-0.62)	77.6(76.5-78.7)	86.9(86.3-87.4)	65.9(64.7-67.0)	92.2(91.8-92.7)
2 claims or 1 discharge in one year	85.3(84.9-85.8)	0.56(0.55-0.57)	55.1(53.8-56.5)	95.2(94.9-95.5)	79.0(77.7-80.3)	86.7(86.2-87.2)
2 claims or 1 discharge in two years	86.8(86.4-87.3)	0.63(0.62-0.65)	68.7(67.5-70.0)	92.7(92.3-93.1)	75.5(74.3-76.7)	90.1(89.6-90.5)
2 claims or 1 discharge in three years	86.7(86.3-87.2)	0.65(0.64-0.66)	74.7(73.5-75.8)	90.7(90.2-91.1)	72.4(71.2-73.5)	91.6(91.2-92.1)
2 claims in 2 years or 1 discharge back to 1988	87.0(86.5-87.4)	0.64(0.63-0.65)	70.6(69.4-71.8)	92.3(91.9-92.7)	75.0(73.9-76.2)	90.6(90.1-91.0)
2 claims in 3 years or 1 discharge back to 1988	86.8(86.4-87.3)	0.65(0.64-0.66)	75.9(74.8-77.1)	90.4(89.9-90.8)	72.0(70.9-73.2)	92.0(91.6-92.4)
1 discharge in three yrs	76.5(76.0-77.1)	0.08(0.07-0.09)	6.4(5.7-7.0)	99.4(99.3-99.6)	79.0(75.2-82.8)	76.5(75.9-77.0)

**Table 3: Accuracy of hypertension case definition algorithms in different patient populations**

<b>Case Definition Algorithm</b>	<b>Percent Agreement</b>	<b>Kappa</b>	<b>Sensitivity</b>	<b>Specificity</b>	<b>Positive Predictive Value</b>	<b>Negative Predictive Value</b>
<b>In outpatients aged older than 65 years (from the primary care chart audit):</b>						
2 claims in one year	66.6(63.7-69.4)	0.39(0.34-0.43)	49.3(45.5-53.1)	95.8(93.9-97.8)	95.3(93.0-97.5)	52.7(49.0-56.4)
2 claims in three years	81.1(78.7-83.5)	0.61(0.57-0.66)	78.4(75.3-81.6)	85.7(82.2-89.2)	90.3(87.9-92.7)	70.1(65.9-74.4)
1 claim or 1 discharge in three years	83.4(81.2-85.7)	0.64(0.59-0.69)	89.4(87.1-98.1)	73.2(68.8-77.7)	85.0(82.3-87.7)	80.3(76.2-84.4)
<b>In patients discharged after hospitalization for myocardial infarction or heart failure (from the EFFECT audit):</b>						
2 claims in one year	60.7(60.0-61.3)	0.18(0.17-0.19)	24.2(23.4-25.1)	92.8(92.3-93.3)	74.7(73.2-76.2)	58.2(57.4-58.9)
2 claims in three years	67.1(66.4-67.7)	0.33(0.33-0.35)	50.8(49.8-51.8)	81.4(80.7-82.1)	70.6(69.6-71.7)	65.3(64.5-66.1)
1 claim or 1 discharge in three years	69.0(68.3-69.6)	0.38(0.37-0.39)	73.7(72.9-74.6)	64.8(63.9-65.7)	64.8(63.9-65.7)	73.7(72.8-74.6)

**Figure 1 Annual population adjusted, age and sex standardized prevalence and incidence of hypertension in adults age 35 and over in Ontario using 2 physician billing claims in 3 years from index claim**



**Figure 2 Annual population adjusted and age standardized prevalence and incidence of hypertension in adults age 35 and over in Ontario, by gender, using 2 physician billing claims in 3 years**

**Figure 2 Annual population adjusted and age standardized prevalence and incidence of hypertension in adults age 35 and over in Ontario, by gender, using 2 physician billing claims in 3 years**

