Validation of a Case Definition to Define Hypertension Using Administrative Data

Hude Quan, Nadia Khan, Brenda R. Hemmelgarn, Karen Tu, Guanmin Chen, Norm Campbell, Michael D. Hill, William A. Ghali, Finlay A. McAlister, for the Hypertension Outcome and Surveillance Team of the Canadian Hypertension Education Programs

Abstract—We validated the accuracy of case definitions for hypertension derived from administrative data across time periods (year 2001 versus 2004) and geographic regions using physician charts. Physician charts were randomly selected in rural and urban areas from Alberta and British Columbia, Canada, during years 2001 and 2004. Physician charts were linked with administrative data through unique personal health number. We reviewed charts of ≈50 randomly selected patients >35 years of age from each clinic within 48 urban and 16 rural family physician clinics to identify physician diagnoses of hypertension during the years 2001 and 2004. The validity indices were estimated for diagnosed hypertension using 3 years of administrative data for the 8 case-definition combinations. Of the 3362 patient charts reviewed, the prevalence of hypertension ranged from 18.8% to 33.3%, depending on the year and region studied. The administrative data hypertension definition of “2 claims within 2 years or 1 hospitalization” had the highest validity relative to the other definitions evaluated (sensitivity 75%, specificity 94%, positive predictive value 81%, negative predictive value 92%, and κ 0.71). After adjustment for age, sex, and comorbid conditions, the sensitivities between regions, years, and provinces were not significantly different, but the positive predictive value varied slightly across geographic regions. These results provide evidence that administrative data can be used as a relatively valid source of data to define cases of hypertension for surveillance and research purposes. (Hypertension. 2009;54:1423-1428.)

Key Words: hypertension ■ surveillance ■ International Disease Classification ■ health information ■ administrative data

Globally, hypertension is one of the most prevalent medical conditions and, in developed nations, one of the most important modifiable risk factors for cardiovascular disease and mortality.1–3 Although treatment of hypertension is associated with a 20% to 25% reduction in cardiovascular events,4–6 it is commonly unrecognized and undertreated. Policy makers and healthcare professionals have made extensive efforts to improve hypertension prevention, detection, and management through national knowledge translation and public health programs.7–8 Population-based hypertension surveillance is a critical tool for evaluating these public health and clinical programs.

Although the ideal surveillance method for identifying hypertension cases and outcomes would be a prospective cohort study with assessment of blood pressure measurements and/or other physiological parameters at repeated intervals, such a method is expensive and may be impractical in many settings. Administrative data are promising potential data sources for surveillance of chronic conditions, because the data are routinely collected, cover wide geographic areas, and have a relatively complete capture of all of the patient encounters with the healthcare system. However, before administrative data can be used for studies and surveillance of hypertension, there is a need to assess the content validity of algorithms used to derive administrative data case definitions of hypertension.

To address this question, Muhajarine et al9 linked Manitoba physician claims data with the Manitoba Heart Health Survey conducted in 1989–1990 to determine the validity of hypertension coding algorithms. They reported that the physician claims data had fair agreement with physical measurement of blood pressure in defining hypertension using 1 claim within 2 years (κ=0.60). Rector et al10 validated physician claims in recording hypertension through self-report survey in the United States in 2000 and reported 72% sensitivity and 80% specificity for the hypertension definition of ≥2 claims within 1 year. Lix et al11 validated Manitoba administrative data using the 2001 Canadian Community Health Survey data.
Compared with self-reported hypertension from survey data, the definition of 2 physicians claims or 1 hospitalization in 3-year administrative data had the highest validity (sensitivity: 76%; positive predictive value [PPV]: 75%); adding prescription medication to the algorithm increased the sensitivity but reduced the PPV. Tu et al\textsuperscript{12} assessed a claim-based case definition for the definition of hypertension using Ontario provincial administrative databases and general practitioner/family physician (GP/FP) chart reviews. The case definition of 2 claims or 1 hospitalization in 3-year administrative data had high sensitivity (78%) and PPV (84%). However, these studies did not assess variation of the validity across region and time period.

We compared the validity of International Classification of Disease (ICD), 9th and 10th versions, case definitions for hypertension using Ontario provincial administrative databases and general practitioner/family physician (GP/FP) chart reviews. The case definition of 2 claims or 1 hospitalization in 3-year administrative data had high sensitivity (78%) and PPV (84%). However, these studies did not assess variation of the validity across region and time period.

We compared the validity of International Classification of Disease (ICD), 9th and 10th versions, case definitions for defining hypertension using Canadian administrative data of hospital discharge abstracts and physician fee-for-service claims between 2 regions (provinces of Alberta and British Columbia, in both rural and urban practices) and 2 separate years (2001 and 2004).

**Methods**

**Recruitment of GPs/FPs and Selection of Patients**

We included fee-for-service GPs/FPs who practiced \( \geq \)2 days per week at their current location between 1999–2001 or 2002–2004 and excluded physicians whose primary practice was at walk-in clinics, community health centers, hospitals, or emergency departments or locum physicians in 2 large cities, Calgary (population \( \approx \)1.1 million) and Vancouver (population \( \approx \)2.1 million), and in 7 rural areas with a population \( < \)10 000 in the provinces of Alberta and British Columbia. A complete list of GP/FPs in these regions was obtained from the Alberta and British Columbia provincial licensing physician directories in Canada. These lists include the GP/FP and clinic name, as well as contact information. Because the contact information on the list is not updated regularly, the GP/FP telephone numbers were verified through the local telephone directories and a Web site search. GP/FPs with a verified telephone number were randomly selected and contacted by fax or telephone to determine their eligibility and to invite them to participate in the study. In addition to the sample, we also included some convenient GP/FP clinics, because the response rate for GPs was low.

Ethics approval was obtained from the University of Calgary and University of British Columbia ethics committees. Participating physicians provided informed consent.

At GP/FP clinics, we used 2 approaches for random selection of patient charts on the basis of the presence or absence of a computerized patient list. For approach 1, where computerized patient information was available, we generated a random list of patients. For approach 2, where computerized patient information was not available, the chart shelves were equally divided into 50 sections. Charts were reviewed consecutively in a section until 1 eligible chart was identified. If there was no eligible chart in a particular section, we replaced this section with another randomly selected section. We identified \( \geq \)50 eligible charts at each clinic for data extraction. The eligibility criteria were those who were \( \geq \)35 years of age, were alive and did not migrate out of the province in the 2-year period before the study years, and had \( \geq \)2 visits to a GP/FP for each period of 1999–2001 for the study population from 2001 and each period of 2002–2004 for the study population from 2004.

**Chart Data Collection and Defining Hypertension**

Five chart reviewers underwent training in the data extraction process through discussion of the definition of data extraction elements and reviewing 10 charts together. A diagnosis of hypertension was defined on the basis of blood pressure readings or a documented diagnosis of hypertension in the GP/FP chart following the Canadian hypertension guidelines (see elsewhere for details).\textsuperscript{13} Pregnancy-induced hypertension was excluded in our definition.

The discrepancies among the reviewers were discussed and resolved. To evaluate agreement among the reviewers, they independently extracted variables from the same 40 charts. The \( k \) value among the reviewers for the presence or absence of hypertension ranged from 0.95 to 1.00, which is considered to be “perfect agreement.”\textsuperscript{14} The reviewers then independently abstracted information from charts, including demographic data, blood pressure readings, medication, and comorbid conditions. We collected information from charts on the presence of 13 comorbidities, including stroke, dementia, diabetes mellitus, dyslipidemia, coronary artery disease, peripheral vascular disease, congestive heart failure, chronic pulmonary disease, asthma, cancer, depression, chronic kidney disease, and dialysis.

**Defining Hypertension Using Administrative Data**

Four administrative databases from Alberta and British Columbia were linked with the chart data using the personal health number as the unique identifier. The databases included the population registry, hospital discharge abstracts, and physician claims from Alberta and British Columbia, as well as emergency department data from Alberta from 1999 to 2004. The population registry contains the demographic information of health care recipients. Canada has a government-financed universal health insurance system; thus, the population registry includes virtually all provincial residents. We used the registry file to determine residence, death, and migration in the study period.

The hospital discharge abstracts include all of the inpatient separations (by discharge or death) from all of the hospitals in the province. Coders reviewed inpatient charts and extract data on personal health number, demographics, diagnoses, procedures, and physician specialty. Before 2001, each discharge record contained \( \leq 16 \) ICD-9 diagnosis codes and, since 2001, each record contained \( \leq 25 \) ICD-10 diagnosis codes.

Fee-for-service physicians in Canada submit claims to their respective provincial government insurance program. The physician claims database contains information on patients' personal health number, physician unique identifier, 1 ICD-9 diagnosis code in British Columbia, and \( \leq 3 \) diagnosis codes and 1 procedure code in Alberta. The physician claims capture nearly all of the outpatient physician services and the majority of the inpatient services.

In the province of Alberta, clinical information from emergency department visits is also collected through the Ambulatory Care Classification System. This database contains \( \leq 10 \) ICD-9-Clinical Modification/ICD-10 diagnosis coding fields in each record. We identified patients with hypertension in the administrative databases using the relevant ICD-9 and ICD-10 codes (ICD-9 codes: 401.x, 402.x, 403.x, 404.x, and 405.x, ICD-10 codes: 110.x, 111.x, 112.x, 113.x, and 115.x) in the \( \geq 25 \) coding fields for diagnosis in the hospital discharge data, \( \leq 3 \) fields in physician claims data, and \( \leq 10 \) fields in emergency department data.

We defined hypertension for each patient using the following 8 case definitions for a combination of hospital discharge and physician claims using 3-year administrative data: (1) 2 physician claims within a 1-, 2-, and 3-year period or 1 hospital discharge for hypertension; (2) 2 physician claims within in a 1-, 2-, and 3-year period for hypertension; (3) 1 physician claim or 1 hospital discharge for hypertension; and (4) 1 physician claim for hypertension. We repeated these definitions in 2-year administrative data. The year period was determined by the time gap between 2 diagnosis claims. For example, when a patient had a physician claim for hypertension, then a second claim for hypertension within the specified time period was required for the patient to meet the case definition for the 2 physician claim algorithms. Presence or absence of hypertension was assigned to each patient 16 times, once for each case definition in 3-year data and then in 2-year data. The first date of the hypertension diagnosis was assigned to patients for case definitions with \( \geq 1 \) hypertension diagnosis. We also tested these case definitions through use of 1 as well as 3 diagnosis coding fields in the physician claims data and by adding emergency department visits.
Statistical Analysis
Descriptive statistics were used for common demographic and comorbid variables. We calculated sensitivity, specificity, PPV, and negative predictive value (NPV) for each case definition, accepting the chart data as the reference standard. The statistic was used to assess agreement between the 2 data sources for the presence of hypertension, considering that chart data may not be accepted as the gold standard. The 95% CI was calculated for these statistics. Considering that validity might vary by age, sex, and 13 comorbidities, we adjusted these factors using log-linear regression and calculated relative sensitivity and PPV between time periods (2004 versus 2001) and regions (urban versus rural area and British Columbia versus Alberta). Relative sensitivity (or PPV) at 1 means that both groups had the same sensitivity (or PPV); <1.0 or >1.0 means that one group had lower or higher sensitivity (or PPV) than another group.

Results
We randomly selected and reviewed 3362 charts at 64 GP/FP clinics to determine the presence of hypertension (Table 1). The prevalence of hypertension ranged from 18.8% to 33.3%, depending on the year, province, or rural versus urban region studied.

The validity of administrative data in determining the presence of hypertension compared with chart data as the reference standard varied across case definitions and length of administrative data observation (Table 2). The validity of the case definitions was relatively stable regardless of the gaps between the 2 claims. For example, sensitivities for the “2-claims-or-1-hospitalization” definition were 76%, 75%, and 74% for gaps among 2 claims within 3, 2, and 1 year in the 3-year administrative data, respectively. However, the validity depended on the length of administrative data observation. For the case definition of 2 claims within 2 years or 1 hospitalization, a 3-year period of administrative data observation had higher sensitivity (75% versus 66%) but lower PPV (81% versus 87%) than a 2-year period of administrative data observation.

Stratification of validity by region and time is shown in Table 3. The sensitivity for the case definition of 2 claims

<table>
<thead>
<tr>
<th>Variable</th>
<th>Province of Alberta</th>
<th></th>
<th>Province of British Columbia</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of general practitioner clinics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients</td>
<td>586</td>
<td>393</td>
<td>586</td>
<td>297</td>
</tr>
<tr>
<td>Age, mean, y</td>
<td>52.6</td>
<td>51.0</td>
<td>53.9</td>
<td>45.5</td>
</tr>
<tr>
<td>Men, %</td>
<td>31.9</td>
<td>41.0</td>
<td>32.9</td>
<td>45.3</td>
</tr>
<tr>
<td>Presence of comorbidity in chart data, %*</td>
<td>46.6</td>
<td>59.8</td>
<td>59.6</td>
<td>43.8</td>
</tr>
</tbody>
</table>

*Comorbidity includes stroke, dementia, diabetes mellitus, dyslipidemia, coronary artery disease, peripheral vascular disease, congestive heart failure, chronic pulmonary disease, asthma, cancer, depression, chronic kidney disease, and dialysis.

Table 1. Characteristics of Study Samples

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of general practitioner clinics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, mean, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence of comorbidity in chart data, %*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Comorbidity includes stroke, dementia, diabetes mellitus, dyslipidemia, coronary artery disease, peripheral vascular disease, congestive heart failure, chronic pulmonary disease, asthma, cancer, depression, chronic kidney disease, and dialysis.

Table 2. Validity of Administrative Data Hypertension Case Definition Compared With Chart Data in 3- and 2-Year Administrative Data

<table>
<thead>
<tr>
<th>Hypertension Definition</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>PPV (95% CI)</th>
<th>NPV (95% CI)</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>PPV (95% CI)</th>
<th>NPV (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 claims or 1 hospitalization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gap between 2 claims</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within 3 y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within 2 y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within 1 y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 claims</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gap between 2 claims</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within 3 y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within 2 y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within 1 y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 claim or 1 hospitalization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Quan et al  Validation and Hypertension  1425

Quan et al  Validation and Hypertension  1425
within 2 years or 1 hospitalization in the 3-year administrative data was higher for year 2001 than that for year 2004 (77% versus 74%), for rural area than that for urban area (79% versus 73%), and for British Columbia than that for Alberta (76% versus 74%). The specificity and NPV were similar across regions and time periods. The validity of the case definition was higher among women than among men, similar across regions and time periods. The validity of the various case definitions assessed was relatively stable across time and geographic regions. However, we found that the presence of comorbidities compared with those with comorbidities, and in those aged ≥65 years relative to those aged <65 years.

To adjust for important confounders, we calculated relative sensitivity and PPV between time periods and regions. For example, this calculated the sensitivity of year 2004 over the sensitivity of year 2001 adjusting for these important confounders. When the relative sensitivity equals 1, administrative data in year 2004 and 2001 have the same sensitivity. When the relative sensitivity is >1, the data in year 2004 have a greater sensitivity compared with data in year 2001.

Table 3. Validity for the Case Definition of 2 Claims Within 2 Years or 1 Hospitalization in 3-Year Administrative Data

<table>
<thead>
<tr>
<th>Year</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>PPV, %</th>
<th>NPV, %</th>
<th>κ</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>77</td>
<td>95</td>
<td>79</td>
<td>94</td>
<td>0.72</td>
</tr>
<tr>
<td>2004</td>
<td>74</td>
<td>94</td>
<td>83</td>
<td>92</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Table 4. Relative Sensitivity (95% CI) and Relative PPV (95% CI) for Hypertension Defined Using 2 Claims Within 2 Years or 1 Hospitalization Case Definition in 3-Year Administrative Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Relative Sensitivity (95% CI)</th>
<th>Relative PPV (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comorbidity (presence vs absence)</td>
<td>0.93 (0.86 to 1.01)</td>
<td>1.11 (1.02 to 1.22)</td>
</tr>
<tr>
<td>Age (≥65 vs &lt;65), y</td>
<td>1.17 (1.08 to 1.27)</td>
<td>1.08 (1.01 to 1.15)</td>
</tr>
<tr>
<td>Sex (female vs male)</td>
<td>0.97 (0.89 to 1.04)</td>
<td>1.01 (0.95 to 1.08)</td>
</tr>
<tr>
<td>Year (2004 vs 2001)</td>
<td>0.98 (0.91 to 1.06)</td>
<td>1.00 (0.94 to 1.07)</td>
</tr>
<tr>
<td>Region (urban vs rural)</td>
<td>0.95 (0.88 to 1.03)</td>
<td>1.08 (1.01 to 1.06)</td>
</tr>
<tr>
<td>Province (BC vs AB)</td>
<td>1.03 (0.96 to 1.12)</td>
<td>0.92 (0.86 to 0.98)</td>
</tr>
</tbody>
</table>

Relative sensitivity and PPV = 1 means that both groups had the same sensitivity/PPV; <1.0/1.0 means that one group had lower/higher sensitivity/PPV than another group. AB indicates province of Alberta; BC, province of British Columbia.

*Comorbidity includes stroke, dementia, diabetes mellitus, dyslipidemia, coronary artery disease, peripheral vascular disease, congestive heart failure, chronic pulmonary disease, asthma, cancer, depression, chronic kidney disease, and dialysis.

95% CI: 0.88 to 1.03). Compared with rural areas, the relative PPV for urban areas was significantly higher (adjusted relative PPV: 1.08; 95% CI: 1.01 to 1.06).

In Alberta, ≥3 ICD-9 diagnosis-coding fields are available for each claim. The majority (74%) of claims had only 1 diagnosis recorded, whereas 4% recorded 2 diagnoses, 2% recorded 3 diagnoses, and 20% were missing diagnoses (services for diagnostics and laboratory tests are not required to record diagnoses). The validity of the various case definitions was not any better when >1 coding field was available (Table 5). Similarly, adding emergency department visits to physician claims and hospital discharge abstracts contributed little to the validity of the case definition.

**Discussion**

We assessed the validity of Canadian inpatient and outpatient administrative data for defining hypertension across time, rural and urban areas, and provinces, as well as patient characteristics. Of the 8 case definitions assessed, the administrative data-based definition of 2 physician claims within 2 years or 1 hospitalization had the best validity. Validity varied slightly across urban and rural areas and provinces. Application of the definition requires ≥3 years of administrative data.

Which case definition should be used for defining hypertension using administrative data? We found that validity of case definitions of 2 claims within 2 years or 1 hospitalization in the 3-year data was higher than that in the 2-year data. However, the definition requires 3 fiscal or calendar years of data in contrast to the 2-year data. Previous studies by Lix et al11 and Tu et al12 demonstrate poor validity using a 1-year length of observation. Considering the results of these previous studies, we recommend 2 claims within 2 years or 1 hospitalization to define hypertension in administrative data using 3-year data. In our study, this definition provides a sensitivity of 75% and a PPV of 81%. Importantly, the validity of this algorithm was relatively stable across time and geographic regions. However, we found that the presence of...
comorbidity was associated with the validity of the algorithms. The sensitivity decreased and PPV increased (Table 4) among patients with comorbidities compared with those without comorbidities. This may be related to the potential that physicians seeing patients with multiple comorbidities are more likely to bill for a comorbid condition and less likely to bill for hypertension.16,17

Is one data source, either physician claims or hospital discharge abstracts, sufficient to define hypertension? Hospital discharge abstracts are available in many countries and have high validity in recording hypertension. Quan et al18 reviewed 4008 in-patient charts and linked the chart data with hospital discharge administrative data in Alberta. They reported 78.6% sensitivity and 91.4% PPV for hypertension recorded in the hospital discharge data. Although the hospital discharge data has high validity in recording hypertension, this database alone is insufficient to determine hypertension prevalence and incidence, because the majority of diagnosis and treatment of hypertension occurs in the outpatient setting. Using physician claims data alone, Lix et al11 were able to identify 97.6% of patients with hypertension in the survey, with the remaining 2.4% of cases obtained from hospital discharge data. In this study, we found that the 2-claims-within-2-years definition had a sensitivity of 73%, specificity of 95%, PPV of 82%, and NPV 91%. Adding hospital discharge data to physician claims data did not increase the validity significantly. Physician claims alone could be used to define prevalence or incidence of physician-diagnosed hypertension but exclusion of inpatient data from the algorithm is likely to miss patients with significant comorbid conditions and could result in poor prediction of mortality or complications. On the basis of the evidence, we recommend defining hypertension using a combination of the hospital discharge abstract and physician claims databases for outcome studies.

This study has limitations. First, chart data as a reference standard only reflect part of the validity of the administrative data. Physician misdiagnoses and incomplete documentation of clinical information in charts were not addressed in this evaluation of administrative data validity. However, chart data are the optimal available data sources for establishing reference standards and are unlikely to bias the study results, particularly given the primary purpose of this study, which was to validate an administrative data case definition for detection of physician diagnosis of hypertension. Second, patients with hypertension who did not visit a GP/FP or were not admitted to the hospital during the study period were missed. Third, we compared time variation across 3 years; changes in data quality may have been observed with longer time periods of observation.

In conclusion, defining hypertension using 2 physician claims within 2 years or 1 hospitalization in 3-year administrative data had substantial validity. The validity varied slightly by geographic region and over time. Applying the definition in 2-year administrative data is likely to miss hypertension. These results provide evidence that administrative data provide a relatively valid source of data to define cases of hypertension for population surveillance and research purposes.

Perspectives

Readily accessible and analyzable population-based administrative data have great potential to facilitate research in hypertension, including surveillance of this common medical condition. The fundamental question, however, before using this information is whether the data are valid for such purposes. Our study addressed this question by comparing administrative and chart data across both time periods and regions in Canada. We found that the administrative data were valid for defining hypertension cases and that the validity varied only slightly by geographic region when the 2 claims within 2 years or 1 hospitalization hypertension case definition was used in administrative data. This definition could be used to define hypertension cases at a population level. Our study also provided evidence to exclude the potential influence of region and time on data validity for determining hypertension incidence and prevalence for surveillance. Although we validated a hypertension case definition in Canadian data, we encourage researchers to undertake similar validation studies to ensure validity of their own regional or national administrative data before using these data.

Acknowledgments

We thank the members of Hypertension Outcome and Surveillance Team of the Canadian Hypertension Education Programs Outcomes Research Task Force, who are listed at http://hypertension.ca/chep/about/committees.

Sources of Funding

This study was funded by Canadian Institutes of Health Research and research grants from Pfizer Canada, Sanofi Aventis, and Merk. H.Q., B.R.H., W.A.G., M.D.H., and F.A.M. are supported by the Alberta Heritage Foundation for Medical Research salary awards. H.Q. and N.K. are supported by a New Investigator Award from the Canadian Institutes of Health Research. W.A.G. is supported by Government of Canada chairs. N.C. is supported...
Disclosures

N.C. is supported by Canadian Chair in Hypertension Prevention and Control. He is the president of Blood Pressure Canada and holds a Blood Pressure Canada grant funded by Servier Canada to develop knowledge translation methods for hypertension control. N.C. has received honoraria payment for hypertension talks and presentations by Bayer, Sanofi, and Bioval. BMS. He also has a consultancy relationship with Novartis, Shering Plough, and Pfizer.

References

Validation of a Case Definition to Define Hypertension Using Administrative Data
Hude Quan, Nadia Khan, Brenda R. Hemmelgarn, Karen Tu, Guanmin Chen, Norm Campbell, Michael D. Hill, William A. Ghali and Finlay A. McAlister
for the Hypertension Outcome and Surveillance Team of the Canadian Hypertension Education Programs

Hypertension. 2009;54:1423-1428; originally published online October 26, 2009;
doi: 10.1161/HYPERTENSIONAHA.109.139279

Hypertension is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2009 American Heart Association, Inc. All rights reserved.
Print ISSN: 0194-911X. Online ISSN: 1524-4563

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://hyper.ahajournals.org/content/54/6/1423

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Hypertension can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Hypertension is online at:
http://hyper.ahajournals.org/subscriptions/