The Accuracy of Administrative Data Diagnoses of Systemic Autoimmune Rheumatic Diseases

SASHA BERNATSKY, TINA LINEHAN, and JOHN G. HANLY

ABSTRACT. ***Objective.*** To examine the validity of case definitions for systemic autoimmune rheumatic diseases [SARD; systemic lupus erythematosus (SLE), systemic sclerosis (SSc), myositis, Sjögren's syndrome, vasculitis, and polymyalgia rheumatica] based on administrative data, compared to rheumatology records.

Methods. A list of rheumatic disease diagnoses was generated from population-based administrative billing and hospitalization databases. Subjects who had been seen by an arthritis center rheumatologist were identified, and the medical records reviewed.

Results. We found that 844 Nova Scotia residents had a diagnosis of one of the rheumatic diseases of interest, based on administrative data, and had had ≥ 1 rheumatology assessment at a provincial arthritis center. Charts were available on 824 subjects, some of whom had been identified in the administrative database with > 1 diagnosis. Thus a total of 1136 diagnoses were available for verification against clinical records. Of the 824 subjects, 680 (83%) had their administrative database diagnoses confirmed on chart review. The majority of subjects who were "false-positive" for a given rheumatic disease on administrative data had a true diagnosis of a similar rheumatic disease. Most sensitivity estimates for specific administrative data-based case definitions were > 90%, although for SSc, the sensitivity was 80.5%. The specificity estimates were also > 90%, except for SLE, where the specificity was 72.5%.

Conclusion. Although health administrative data may be a valid resource, there are potential problems regarding the specificity and sensitivity of case definitions, which should be kept in mind for future studies. (J Rheumatol First Release May 1 2011; doi:10.3899/jrheum.101149)

Key Indexing Terms:
ADMINISTRATIVE DATABASE — VALIDATION
RHEUMATOLOGY — SYSTEMIC AUTOIMMUNE RHEUMATIC DISEASES

Chronic disease surveillance is important to clinicians, patient advocacy groups, researchers, and health policy makers. Administrative healthcare databases have been used for surveillance, outcomes research, and quality assessment in chronic rheumatic disease[1]. However, there have been very few validation studies of these data sources. We examined the accuracy of case definitions of systemic autoimmune rheumatic disease (SARD) based on administrative data from the province of Nova Scotia, Canada, by comparing the data to that found in rheumatology clinic records. Six diseases were selected for study: systemic lupus erythematosus (SLE), systemic sclerosis (SSc), myositis, Sjögren's syndrome (SS), systemic vasculitis, and polymyalgia rheumatica (PMR). These were chosen because they are complex rheumatic diseases, for which there are few data regarding the validity of using administrative databases for case ascertainment in the context of surveillance or research[2].

From the Division of Clinical Epidemiology and the Division of Rheumatology, Department of Medicine, McGill University Health Centre (MUHC), Montreal, Quebec; and the Division of Rheumatology, Department of Medicine, and the Department of Pathology, Dalhousie University and Queen Elizabeth II Health Sciences Centre, Halifax, Nova Scotia, Canada.

Supported by an operational grant from the Canadian Institutes of Health Research (CIHR). S. Bernatsky is a CIHR Junior Investigator and Canadian Arthritis Network Scholar and is supported by the McGill University Department of Medicine and Research Institute. Additional financial and research support from the Public Health Agency of Canada.

S. Bernatsky, MD, PhD, Division of Clinical Epidemiology and the Division of Rheumatology, Department of Medicine, MUHC; T. Linehan, MPA, Division of Rheumatology, Department of Medicine, Dalhousie University and Queen Elizabeth II Health Sciences Centre; J.G. Hanly, MD, Division of Rheumatology, Department of Medicine, and the Department of Pathology, Dalhousie University and Queen Elizabeth II Health Sciences Centre.

Address correspondence to Dr. S. Bernatsky, Montreal General Hospital, Clinical Epidemiology, 1650 Cedar Avenue, L10-424, Montreal, Quebec H3G 1A4, Canada. E-mail: sasha.bernatsky@mail.mcgill.ca

Accepted for publication March 4, 2011.

MATERIALS AND METHODS

Nova Scotia is a Canadian province of about 950,000 residents, with universal coverage for health services (including physician services and hospitalizations). Information on all physician billing activity (including diagnosis) and hospitalizations (including the discharge diagnoses) is maintained by the Population Health Research Unit (PHRU) in the Department of Community Health and Epidemiology at Dalhousie University, Halifax. In each province in Canada, physician services are comprehensively recorded for remuneration purposes; in most instances, a single diagnostic code is provided for each contact (as indicated by the physician on the billing claim), for both outpatient and inpatient services. For hospitalization records, a primary diagnosis is available, as well as multiple nonprimary diagnoses; these are extracted by trained medical clerks according to defined protocols.

For our study, after obtaining ethics approval, we obtained from PHRU a list of Nova Scotia residents who had been identified as having ≥ 1 of sever-

al rheumatic diseases as of 2005, based on billing and/or hospitalization data collected since 1990. To identify these individuals, the case definition was based on International Classification of Diseases, 9th ed (ICD-9) diagnostic codes (Table 1). Cases were identified in 3 ways using billing data (2 different algorithms) or hospitalization data. The first algorithm required ≥ 2 billing claims by any physician with the relevant ICD-9 code, provided the 2 billing events were at least 2 months apart. The second algorithm required > 1 claim by a rheumatologist or internist. For the hospitalization data, cases were defined based on the basis of ≥ 1 discharge diagnosis for the particular diagnosis as per ICD-9 codes. Some, but not all, of the cases were identified by all 3 case definitions; some had been diagnosed by ≥ 2 of the case definitions; and some had been diagnosed by only 1 case definition.

To assess the validity of the algorithms used to identify rheumatic disease cases in the PHRU administrative databases, we relied on the clinical diagnoses made by 8 rheumatologists at The Arthritis Centre of Nova Scotia, Division of Rheumatology, Queen Elizabeth II Health Sciences Centre, Halifax, serving the Capital Health Region. (There are about 2700 physicians practicing in Nova Scotia, with a third of these being in the Halifax area, where the Arthritis Centre is located.) A case of a specific rheumatic disease, identified within the administrative data, was considered a "true-positive" if the same individual also had clinical evidence of that disease recorded as the final diagnosis by ≥ 1 of the rheumatologists. True-negatives were those subjects (among the cases of other rheumatic diseases identified by PHRU) who were not identified with a given rheumatic disease in the administrative data, and who indeed did not have that specific rheumatic disease on chart review. For each specific rheumatic disease, sensitivity was determined (the ratio of true positives vs the number of true cases overall) as well as specificity (the ratio of the true negatives vs the number of individuals who truly do not have the disease). For these definitions, the assumption was that the true diagnosis was the final rheumatology diagnosis on the clinic records.

RESULTS

We found that 844 Nova Scotia residents had a diagnosis of one of the rheumatic diseases of interest, based on administrative data, and had had ≥ 1 rheumatology assessment at The Arthritis Centre. Twenty of these were excluded because no charts were available, leaving 824 subjects for our analyses. Some of the 824 individuals had > 1 of the rheumatic disease diagnoses identified in the administrative database, so that there were a total of 1136 diagnoses from administrative data that could be validated by comparison with the clinical records.

Of the 1136 administrative diagnoses, 110 were nonspecific ICD-9 codes (710.8 — other specified diffuse diseases of connective tissue, and 710.9 — unspecified diffuse connective tissue disease). Among the remaining 1026 administrative diagnoses, in 824 individuals, 680 individuals were correctly classified according to our review of the clinical records. Most of the false-positives for 1 SARD had a "true diagnosis" of another similar rheumatic disease. For example, in the 165 subjects who were false-positive for SLE by administrative data (that is, not confirmed by chart review), in most cases the true diagnosis (as recorded in the medical chart) was another connective tissue disease (in order of frequency the diagnoses were SSc, primary SS, mixed connective tissue disease, undifferentiated connective tissue disease, inflammatory myositis, limited cutaneous systemic sclerosis, inflammatory arthritis, fibromyalgia, antiphospholipid syndrome, and osteoarthritis).

Table 1 indicates the results for each rheumatic disease, including true-positive diagnoses, false-positive diagnoses, sensitivity, and specificity, calculated using the total pool of 1136 diagnoses (in 824 individuals). The sensitivity of the administrative data was relatively low for SSc (as compared to the other rheumatic diseases) at 80.5%, while the administrative data had comparatively low specificity for SLE (72.5%). Of the 386 SLE cases identified in the administrative data, 165 were false-positives according to our strict definition (that the final clinical diagnosis was the gold standard). The SLE false-positive rate was 42.7%, the highest false-positive rate

Table 1. Results for systemic autoimmune rheumatic disease diagnoses (n = 1136) identified from administrative data: true-positives (using chart review as the gold standard), false-positives, sensitivity, and specificity.

Disease* (ICD-9 code): 1136 Diagnoses in 824 Individuals	Administrative Diagnosis (All Positive)	Diagnosis on Chart Review	True-Positive	False-Positive	True-Negative	False-Negative	Sensitivity, % (95% CI)	Specificity, % (95% CI)
SLE (710.0)	386	225	221	165	434	4	98.2 (95.5, 99.3)	72.5 (68.7, 75.9)
SSc (710.1)	104	82	66	38	704	16	80.5 (70.6, 87.6)	94.9 (93.0, 96.2)
Myositis (710.3–4)	66	43	38	28	753	5	88.4 (75.5, 94.9)	96.4 (94.9, 97.5)
Sjögren's syndrome (710.2)	115	88	84	31	705	4	95.5 (88.8, 98.2)	95.8 (94.1, 97.0)
Vasculitis (446)	120	92	86	34	698	6	93.5 (86.5, 97.0)	95.4 (93.6, 96.7)
PMR (725)	235	186	185	50	588	1	99.5 (97.0, 99.9)	92.2 (89.8, 94.0)
Other†	110	65	—	—	—	—	—	—

* Cases identified on the basis of hospitalization (≥ 1 discharge diagnosis) or physician billing (> 2 diagnoses by any physician ≥ 8 weeks apart but within 2 years, or ≥ 1 diagnosis by a rheumatologist). † Includes 710.8 — other specified diffuse diseases of connective tissue, and 710.9 — unspecified diffuse connective tissue disease. ICD-9: International Classification of Diseases, 9th ed.; SLE: systemic lupus erythematosus; SSc: systemic sclerosis; myositis: includes dermatomyositis and polymyositis; PMR: polymyalgia rheumatica.

of all the rheumatic conditions. The lowest false-positive rate was in PMR (50/235 = 21.3%).

Using SLE as an example, Figure 1 displays variations in the true-positive rates according to different sources of administrative data (using chart review as the gold standard). There were trends for higher percentages of true-positives for cases identified in the hospital databases (vs those that were not), as well as for those positive both in hospital and in physician billing data (vs those found in only 1 source) and for cases diagnosed by a rheumatologist or internist (vs those diagnosed by another type of physician). However, the confidence intervals of these estimates are wide and overlapping. Similar results were apparent for the other rheumatic diseases.

DISCUSSION

In the developed world, there is increasing interest in monitoring the prevalence of rheumatic disease in our aging populations[3], particularly because the pool of key medical subspecialists (e.g., rheumatologists) is insufficient to meet the clinical need[4,5]. Accurate estimates of rates of disease are critical to determining the clinical significance of these disorders and their effect on public health. Much of the existing literature on outcomes in patients with rheumatic disease is based on data from tertiary care centers. These data are limited by sample size and the potential for selection bias. An alternative approach is to create large population-based cohorts from administrative databases, but this is also subject to bias since the optimal methodology for identifying patients with rheumatic disease is not known. Previous efforts have used a combined approach for case ascertainment, for example by including patient registries and outpatient rheumatology clinics[6], or by using administrative data such as physician billing and hospitalization records[7].

Of the 824 individuals in our study (with 1136 administrative diagnoses), 680 (83%) were given correct diagnoses according to the clinical records. Thus, the vast majority of the individuals who were studied at our center did have evidence of their true underlying diagnosis in the administrative database according to the definitions used here. This suggests that most SARD, at least among those seen by a rheumatologist, are indeed detected within administrative datasets. However, at a truly population-based level, there would undoubtedly be cases not detected in our study, because only cases with the ICD-9 codes and referred to The Arthritis Centre of Nova Scotia were selected for validation. This method might miss cases with milder disease, including those in remission (who may not be seeking care at all) or those with very serious comorbid conditions (e.g., endstage renal disease in SLE), who may be under the care of other specialists. At the same time, among the potential cases identified by our administrative data algorithms, there were a number of false-positive cases.

If one considers the maximal potential effect of false-positive cases on prevalence estimates in SLE, one may anticipate that the prevalence would be inflated by up to 43% (for example, instead of a prevalence of 1 in 2000, the prevalence might be estimated at 1 in 1500). This assumes that there are no false-negatives in our approach, which is generally not the case.

There are few other studies with which to compare our results. Katz, *et al* estimated a sensitivity of 85% when comparing billing code data for SLE versus clinical records, but

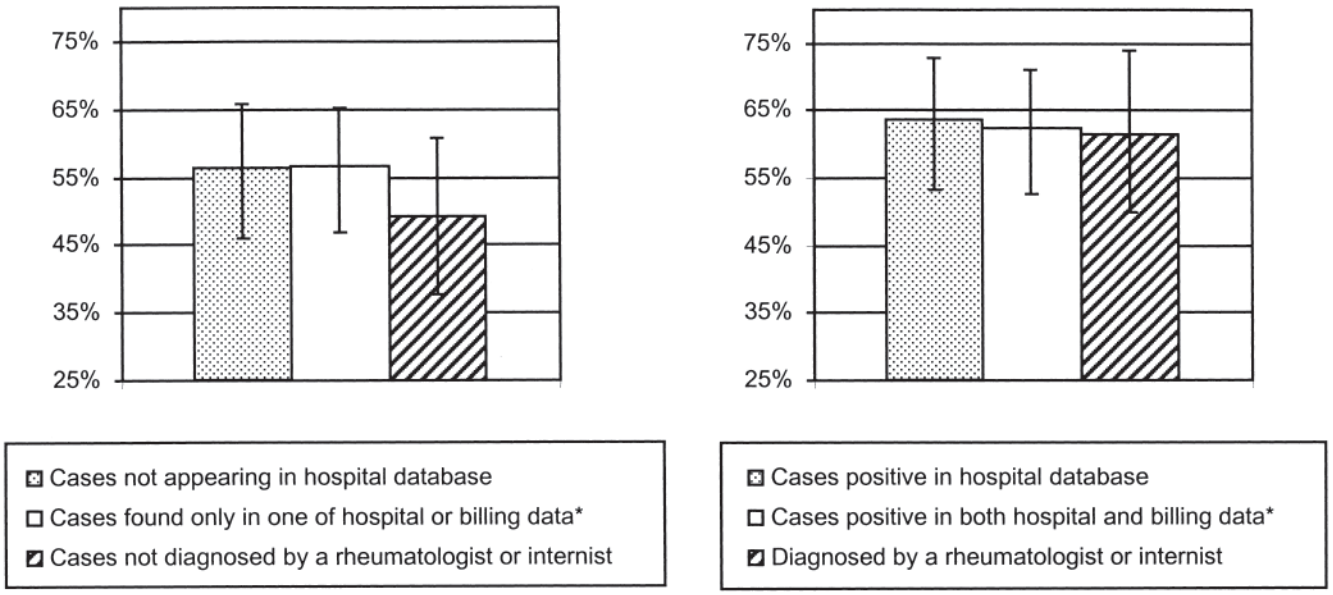

Figure 1. True-positive rates for cases of systemic lupus erythematosus identified from different sources of administrative data (using chart review as the gold standard). *Cases identified on the basis of hospitalization (≥ 1 discharge diagnosis) or physician billing (> 2 diagnoses by any physician ≥ 8 weeks apart but within 2 years), or ≥ 1 diagnosis by a rheumatologist or internist.

that was based only on the billing data of rheumatologists[8]. Lim, *et al* found that the sensitivity and specificity of hospital discharge ICD-9 codes for SLE were 79.3% and 90.2%, respectively, for a 5-year period (with lower sensitivity of 67% and higher specificity of 93.5% for a 1-year period)[9]. A smaller validation study in SSc found that case definitions from hospitalization data were difficult to verify from available clinical records at a single center; the authors reported verification of the ICD-9 diagnostic code for SSc to be only 55.4%[10]. However, the latter findings were most likely due to inadequate documentation in medical charts, which made it difficult to use the American College of Rheumatology classification criteria for SSc (as opposed to simply confirming a clinical diagnosis).

There was a tendency for relatively low specificities for some of our case definitions, particularly SLE. This is partly due to the methodology used for case ascertainment. We defined a case to be a true-positive only when the administrative data diagnosis agreed with the final clinical diagnosis. This meant, for example, that a patient initially thought to have one disease would have the appropriate indication in administrative data. Later, however, the illness may have evolved, leading to a revision of the final diagnosis in the rheumatology chart.

For pharmacoepidemiology research and other studies using administrative data, using at least 2 billing diagnoses and/or at least 1 hospital diagnosis for the disease of interest is becoming the most frequently used methodology. However, researchers who rely on these data sources should additionally consider other strategies to improve specificity of the case definition, such as requiring at least 1 prescription for a medication that would normally be used to treat that disease. For example, in the case of SARD, this might be systemic corticosteroids or other immunomodulators. This approach is only possible when population-based drug prescription data are available (which is not the case for Nova Scotia).

There were trends for higher percentages of true-positives for cases identified in the hospital databases (vs those that were not), as well as for those identified in both hospital and billing data (vs those found in only 1 source) and for cases diagnosed by a rheumatologist or internist (vs those diagnosed by another type of physician). This might support the argument that more stringent criteria (such as limiting the billing data to that of rheumatologists and internists) may improve specificity; this will be at the expense of sensitivity. Studies show that only half of the cases of SARD in Nova Scotia are diagnosed by rheumatologists[11].

We have estimated the sensitivity and specificity of different data sources and approaches using Canadian administrative data across several provinces[11], using Bayesian latent class regression to combine divergent results (e.g., individuals who are defined as a case by ≥ 1 definition but not by all). With this larger dataset, we were able to demonstrate that the properties of case definition algorithms may actually vary according to patient characteristics. In those analyses, the sensitivity of case definitions-based billing data was at best between 70% and 90%, and lowest (50%–70%) in older individuals. For rheumatology billing data, the sensitivity estimates were about 50%–70% in younger individuals and somewhat lower (40%–50%) in older individuals. Rheumatology billing data sensitivity estimates tended to be higher for urban (vs rural) residents. Hospitalization data were the least sensitive.

We note that the billing code diagnoses are potentially dependent on the "gold standard" disease diagnosis in those patients billed by the same rheumatologists who cared for the patients. To minimize verification bias, we tried to make the review of the medical charts for clinical diagnosis confirmation as independent as possible.

It should also be pointed out that the algorithms used to ascertain potential cases in our study are not necessarily applicable to places where different health systems are in place (e.g., not universal), or where physicians may have different skills in using ICD-9 codes. Also, we did not include a review of rheumatoid arthritis billing codes, so we cannot comment on the validity related to this diagnostic code. However, only a tiny fraction of the SARD cases had a final diagnosis of rheumatoid arthritis.

Despite the limitations of our study, the results suggest that health administrative data may be a valid source from which to identify cases of SARD for surveillance and outcomes research. However, our work indicates potential problems regarding the specificity and sensitivity of case definitions for particular diseases. These should be kept in mind for future studies using administrative data on SARD.

ACKNOWLEDGMENT

We are grateful for the support of the attending staff at Dalhousie University and Queen Elizabeth II Health Sciences Centre for allowing us access to their data and records.

REFERENCES

1. Ward MM. Severity of illness in patients with systemic lupus erythematosus hospitalized at academic medical centers. J Rheumatol 2005;32:27-33.
2. Bernatsky S, Pineau C. Administrative database research in SLE: a review of potential usefulness and limitations. In: Ulrich CM, Bellinger KA, editors. Systemic lupus erythematosus research developments. Hauppauge, NY: Nova Science Publishers; 2007.
3. Lix LM, Yogendran MS, Shaw SY, Burchill C, Metge C, Bond R. Population-based data sources for chronic disease surveillance. Chronic Dis Can 2008;29:31-8.
4. Hanly JG. Manpower in Canadian academic rheumatology units: current status and future trends. Canadian Council of Academic Rheumatologists. J Rheumatol 2001;28:1944-51.
5. Hanly JG. Physician resources and postgraduate training in Canadian academic rheumatology centers: a 5-year prospective study. J Rheumatol 2004;31:1200-5.
6. Jonsson H, Nived O, Sturfelt G, Silman A. Estimating the incidence of systemic lupus erythematosus in a defined population using multiple sources of retrieval. Br J Rheumatol 1990;29:185-8.
7. Ward MM, Pajevic S, Dreyfuss J, Malley JD. Short-term prediction

of mortality in patients with systemic lupus erythematosus: classification of outcomes using random forests. Arthritis Rheum 2006;55:74-80.
8. Katz JN, Barrett J, Liang MH, Bacon AM, Kaplan H, Kieval RI, et al. Sensitivity and positive predictive value of Medicare Part B physician claims for rheumatologic diagnoses and procedures. Arthritis Rheum 1997;40:1594-600.
9. Lim SS, Jamal A, Bayakly R, Tong L, Drenkard C. Georgia Lupus Registry — accuracy of hospital discharge data in identifying systemic lupus erythematosus [abstract]. Arthritis Rheum 2007;54 Suppl:S505.
10. Khurma V, Furst DE, Krishnan E, Khanna D. Verification of ICD-CM-9 coding for the diagnosis of systemic sclerosis [abstract]. Arthritis Rheum 2006;54 Suppl:S344.
11. Bernatsky S, Lix L, Hanly JG, Hudson M, Badley E, Peschken C, et al. Surveillance of systemic autoimmune rheumatic diseases using administrative data. Rheumatol Int 2010 Jul 28 (E-pub ahead of print).