Arthritis & Rheumatism (Arthritis Care & Research)
Vol. 61, No. 9, September 15, 2009, pp 1264–1267
DOI 10.1002/art.24793

CONTRIBUTIONS FROM THE FIELD

Polymyalgia Rheumatica Prevalence in a Population-Based Sample

S. BERNATSKY,[1] L. JOSEPH,[2] C. A. PINEAU,[1] P. BELISLE,[1] L. LIX,[3] D. BANERJEE,[1] AND A. E. CLARKE[1]

Objective. **To determine polymyalgia rheumatica (PMR) prevalence using population-based administrative data, and to estimate the error associated with case ascertainment approaches when using these databases.**

Methods. **Cases were ascertained using physician billing and hospitalization data from the province of Manitoba (population 1.1 million). Focusing on the population age ≥45 years, we compared 3 different case definition algorithms and also used statistical methods that accounted for imperfect case ascertainment to estimate the prevalence and the properties of the ascertainment algorithms. A hierarchical Bayesian latent class regression model was developed that also allowed us to assess differences across patient demographics (sex and region of residence).**

Results. **Using methods that account for the imperfect nature of both billing and hospitalization databases, we estimated the prevalence of PMR in women age ≥45 years to be lower in urban areas (754.5 cases/100,000; 95% credible interval [95% CrI] 674.1–850.3) compared with rural areas (1,004 cases/100,000; 95% CrI 886.3–1,143). This regional trend was also seen in men age ≥45 years, where the prevalence was estimated at 273.6 cases/100,000 (95% CrI 219.8–347.6) in urban areas and 380.7 cases/100,000 (95% CrI 311.3–468.1) in rural areas. Billing data appeared more sensitive in ascertaining cases than hospitalization data, and a large proportion of diagnoses was made by physicians other than rheumatologists.**

Conclusion. **These data suggest a higher prevalence of PMR in rural versus urban regions. Our approach demonstrates the usefulness of methods that adjust for the imperfect nature of multiple information sources, which also allow for estimation of the sensitivity of different case ascertainment approaches.**

Introduction

Polymyalgia rheumatica (PMR) is an inflammatory rheumatic disease characterized by diffuse pain and stiffness, particularly in the neck, shoulders, and proximal extremities. It most commonly occurs in older individuals, and is thought to be at least twice as common in women than in men (1). Its etiology is unknown, although genetic factors are believed to contribute (2); one hypothesis is that heightened immunologic responses to environmental insults (e.g., viral infections) may trigger the disease.

Our knowledge of the epidemiology of PMR is largely based on only a handful of studies, with the most recent data only extending to 1999 (3,4). This is due in part to the difficulty in accurately diagnostically capturing cases. Our aim was to estimate the prevalence of PMR in a large general population age ≥45 years using administrative data from 1995 to 2006. We relied on both physician billing and hospitalization databases, and we employed statistical methods that adjusted for possible misclassification within the data sources. In addition, we assessed the sensitivity of different case ascertainment approaches. Our research was approved by the McGill University Ethics Review Board.

Methods

We ascertained cases of PMR using the physician billing and hospitalization databases covering all of the residents of the province of Manitoba (~1.1 million individuals) for the period 1995–2006. The billing database documents

Supported by the Canadian Institutes of Health Research. Dr. Bernatsky is a Canadian Arthritis Network Scholar and was supported by the Canadian Institutes of Health Research, the Fonds de la Recherche en Santé du Québec, and the Department of Medicine, Research Institute, McGill University Health Centre. Drs. Joseph and Clarke are Fonds de la Recherche en Santé du Québec National Scholars. Dr. Pineau's work was supported by the Department of Medicine, Research Institute, McGill University Health Centre.

[1]S. Bernatsky, MD, PhD, C. A. Pineau, MD, P. Belisle, MSc, D. Banerjee, BSc, MDCM, A. E. Clarke, MD, MSc: McGill University Health Centre, Montreal, Quebec, Canada; [2]L. Joseph, PhD: McGill University Health Centre and McGill University, Montreal, Quebec, Canada; [3]L. Lix, PhD: University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

Address correspondence to S. Bernatsky, MD, PhD, Division of Clinical Epidemiology, Research Institute, McGill University Health Centre, 687 Pine Avenue West, V-Building, V2.09, Montreal, Quebec, H3A 1A1, Canada. E-mail: sasha.bernatsky@mail.mcgill.ca.

Submitted for publication November 19, 2008; accepted in revised form June 11, 2009.

physician services for all provincial beneficiaries; only 1 diagnostic code is allowed per visit.

For the hospitalization data, we defined a case as any hospitalization indicating a discharge diagnosis of PMR (primary or nonprimary). In the billing data, cases were first defined according to an algorithm requiring ≥2 physician visits for PMR <2 months apart, but within a 2-year span. In a second alternative algorithm, we defined cases as those where there was ≥1 PMR billing code contributed by a rheumatologist. All Manitoban citizens seeking health care were captured in billing data and those with an inpatient stay were captured in the hospitalization database. However, since our 3 diagnosis definitions differed substantively from each other, some individuals were detected by one definition, but not by another.

The hospitalization database records indicate a primary diagnosis and multiple nonprimary diagnoses for each hospitalization (in 2003 to 2004 and earlier, the data include up to 16 diagnoses per admission, and from 2004 to 2005 onward, the data include up to 25 diagnoses). These diagnoses are abstracted by medical records clerks and are not necessarily the same diagnoses as those contained in the billing database. In both databases, PMR diagnoses are captured as International Classification of Diseases, Ninth Revision code 725. Because PMR is considered to be a disease of persons of late middle age or older, we restricted our analyses to the population of Manitoba that was age ≥45 years.

We generated naive prevalence estimates under the assumption of no error (perfect sensitivity and specificity), applying the 3 case definition algorithms (i.e., 2 for billing data and 1 for hospitalization data) separately. In each case, we divided the number of identified cases by the appropriate population denominator (obtained from Statistics Canada). No confidence intervals are provided for the naive prevalence estimates because the estimates are based on the entire population of the province of Manitoba (not on a sample), assuming no ascertainment error within a given method.

We then generated a fourth prevalence estimate, adjusted for the imperfect sensitivity and specificity of both billing and hospitalization data, using a previously developed Bayesian latent class model that does not assume any data source to be a gold standard (5). It is necessary to consider the sensitivity and specificity of case ascertainment methods because any method of case ascertainment contains some error, although this is often overlooked in analyses using administrative data. All prevalence estimates were calculated for December 31, 2006, based on the number of cases that had been identified during the study period (1995–2006) who remained alive as of December 31, 2006.

Bayesian statistical methods are based on the central Bayes theorem; an unknown parameter (e.g., disease prevalence) is estimated by combining existing information (e.g., published data or expert opinion) with new data. The prior distribution representing preexisting knowledge can be informative (i.e., contain substantive knowledge about the unknown parameter values) or it can contain low information (e.g., a flat probability distribution, where all values for an unknown parameter are equally probable a priori). The results from Bayesian estimates vary according to the prior distributions used; with prior low information distributions, the final estimates depend mostly on the data provided. We produced 95% credible intervals (95% CrIs) for all of our Bayesian estimates, representing the values between which there is a 95% probability of containing the parameter of interest, given the prior information used and the data at hand.

Without access to a gold standard means of case ascertainment, the true sensitivity and specificity of a single diagnostic approach or data source cannot be directly determined; neither can disease prevalence be observed (6). The true disease state for each subject is unknown, or latent. Nevertheless, each method of ascertainment provides some information about the case status of each subject, allowing true disease prevalence to be probabilistically estimated. Within any model, the methods of case ascertainment can be treated as conditionally dependent or independent. Conditional independence exists when, given a subject's true disease status (which is unknown), the result from one method of case ascertainment is statistically independent from the results of the other methods. In our situation, 2 of the ascertainment methods were derived from a similar source (billing claims); therefore, conditional independence was not a plausible assumption. Therefore, in our model, we estimated the dependence between these tests, and adjusted our calculations for this dependence.

With an approach based on 3 tests or data sources, in the face of possible dependence, the statistical problem becomes nonidentifiable. In practice, this means that in order to estimate all of the quantities of interest, one must incorporate informative prior distributions over some of the parameters (7). To do this, we relied on the results from earlier studies of systemic autoimmune rheumatic diseases using administrative databases (8,9), where the specificities of all of the methods of case ascertainment were consistently very high. Therefore, for our primary analyses, we set informative prior distributions for our 2 billing data case ascertainment approaches, corresponding to specificities of 98% (prior 95% CrI 96–100%). We set alternative specificity priors corresponding to specificities of 99% (95% CrI 98–100%) and 94% (95% CrI 88–100%). Because the results using the different sets of priors were all very similar, we only reported results from the first set of priors. For all of the other parameters (including prevalence and sensitivity), we employed very diffuse (low information) prior distributions.

A hierarchical Bayesian latent class regression model was used to estimate disease prevalence and the sensitivities of case ascertainment methods. Our general methods in this context have been previously described in detail (5), but briefly, the levels of our hierarchical model accounted for 1) population sampling variability (assumed to follow a binomial distribution) and misclassification error, adjusting for both false-positives and false-negatives, 2) variation in disease prevalence related to patient demographics (age, sex, and rural versus urban residence), input as a logistic regression model on the binomial probabilities from the first level of the model, and 3) variation in case ascertainment sensitivity according to patient

demographics (age, sex, and rural versus urban residence), input as a distinct parameter for the sensitivity of each case ascertainment approach, with different values for each combination of sex and region. Conditional dependence between the 2 ascertainment methods based on billing tests was handled by the addition of a covariance term, similar to the fixed-effects model by Dendukuri and Joseph in 2001 (7), based on ideas from Vacek (10). WinBUGS statistical software, version 1.4.3 (Medical Research Council, London, UK) was used for our analyses.

Results

Figure 1A displays the prevalence estimates as of December 31, 2006, derived from the primary analyses using the hierarchical Bayesian latent class model that adjusts for the imperfect nature of the databases. The prevalence was higher for women than for men (Figure 1A), and estimates appeared to be lower in urban versus rural areas. In women age ≥45 years, the prevalence was 754.5 cases per 100,000 in urban regions (95% CrI 674.1–850.3) and 1,004 cases per 100,000 in rural areas (95% CrI 886.3–1,143). In men age >45 years, the prevalence was 273.6 cases per 100,000 in urban regions (95% CrI 219.8–347.6) and 380.7 cases per 100,000 in rural areas (95% CrI 311.1–468.1). In contrast, the crude PMR prevalence estimates without the adjustment for imperfect sensitivity and specificity were 587.3 cases per 100,000 for women age ≥45 years and 218.0 cases per 100,000 in men age ≥45 years.

The sensitivity of case ascertainment for PMR using hospitalization data was very low (Figure 1B), ranging from a low of 13.4% (95% CrI 7.0–14.0) in men living in rural areas to a high of 23.6% (95% CrI 17.5–30.3) in men living in urban areas. For billing data, the sensitivity was higher for the algorithm requiring ≥2 physician visits for PMR compared with the algorithm requiring ≥1 rheumatologist diagnosis for PMR. The sensitivity estimates for 2 billing code diagnoses ranged from 67.6% (95% CrI 52.8–81.0) in men living in urban areas to 63.4% (95% CrI 56.5–70.5) in women living in urban areas.

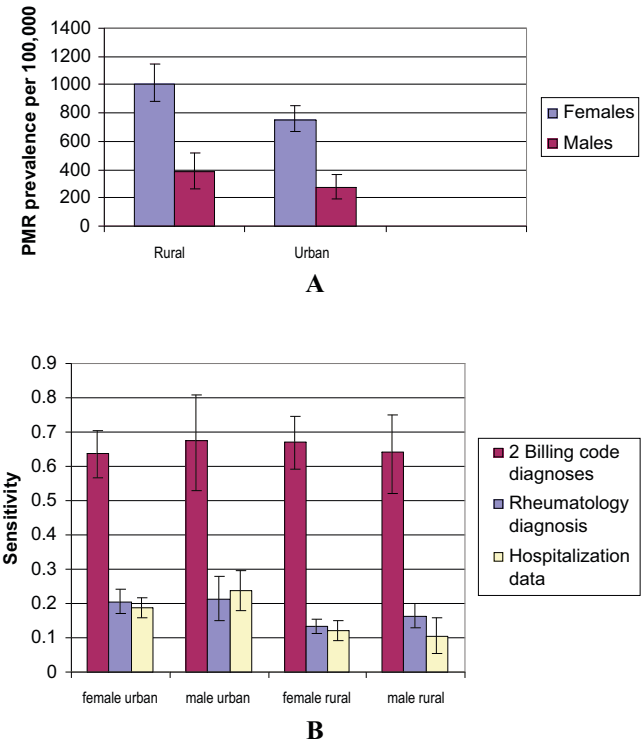

Figure 1. Effects of demographics (sex and rural versus urban residence) on polymyalgia rheumatica (PMR) prevalence and case ascertainment sensitivity by hierarchical Bayesian latent class modeling. **A**, prevalence among the population age ≥45 years, **B**, case ascertainment sensitivity. Error bars show the Bayesian credible intervals, the values between which there is a 95% probability of containing the parameter of interest, given the data and prior information input. Case ascertainment was based on any one of the following algorithms: ≥2 physician claims diagnostic codes for PMR (≥8 weeks apart and within 2 years) contributed by any physician, ≥1 diagnostic PMR code contributed by a rheumatologist, or ≥1 hospitalization discharge diagnostic code (primary or nonprimary) for PMR.

Discussion

The diagnosis of PMR is often a challenge, given that it may present with nonspecific symptoms (including fatigue and other systemic symptoms), generally without any obvious physical findings. Sometimes, patients can even have inflammatory arthritis that resembles rheumatoid arthritis, and authors have noted the absence of clinical or laboratory features that definitively differentiate between these 2 conditions (11). In previous work using administrative data, we showed that case ascertainment of systemic lupus erythematosus may vary according to the training and experience of physicians (9), and this is likely true in PMR as well.

There are very few data regarding the prevalence of PMR. The only other population-based data in North America is from Olmsted County, Minnesota, where the prevalence of PMR was derived from cumulative incidence rates (1,4). Here, the prevalence of PMR as of the year 2000 was estimated at 739 cases per 100,000 people age ≥50 years (95% confidence interval 674–808). No examination of differences in urban versus rural populations was provided; however, a similar female predominance for PMR was demonstrated in Olmsted County (ratio of 1 man to 1.7 women) compared with our data (ratio of 1 man to 2.6 women). Because the risk of PMR increases with age, the overall PMR prevalence may be currently underestimated, given the trend toward aging populations in developed countries.

In terms of comparing our results with the Olmsted County figure, our methods produced a PMR prevalence in Manitobans age ≥45 years (not ≥50 years) because population figures for our denominator stratified by age, sex, and urban versus rural residence are available for residents ages <45 and ≥45 years. However, considering that of Manitoba residents age >45 years, approximately 20% are between ages 45 and 49 years, we would estimate that the prevalence of PMR in Manitobans age ≥50 years would be approximately 641.5 cases per 100,000 in urban areas and 864.2 cases per 100,000 in rural areas. These are compatible with the Olmsted County data.

We acknowledge that diagnoses obtained from administrative databases are not always complete or accurate. For example, the existence of only one diagnostic code per

visit for most physician billing databases in Canada limits the sensitivity of this source because patients (particularly those who are older) may have multiple comorbidities, which may take precedence as the billing diagnosis. Hospitalization data also contain some error, and alone appeared to be a very insensitive means of case ascertainment for PMR (possibly because PMR symptoms would rarely require an inpatient stay). Although one method is not necessarily better than another for case ascertainment, we believe that our work underlines the utility of an approach that uses more than one data source, adjusting for the error in each source.

Our observation of higher PMR prevalence in rural versus urban areas (even after adjustment for age and sex distributions) is interesting, particularly because one might expect that once an individual is diagnosed with a chronic illness, there might be a tendency to gravitate toward urban areas, where more resources (physicians, allied health care, etc.) are concentrated. An increased risk of autoimmune diseases has been associated with farming; specific occupational exposures may include animals, pesticides, organic (e.g., grain) dusts, and crystalline silica, all of which have been suggested to predispose to various systemic autoimmune rheumatic diseases such as rheumatoid arthritis, systemic lupus erythematosus, and systemic sclerosis (12–15).

To date, studies of a possible infectious etiology for PMR have focused on viruses that are characteristically spread between humans, such as hepatitis B, parainfluenza viruses, respiratory syncytial virus, measles virus, herpesviruses, and the Epstein-Barr virus. Data linking these viruses to PMR onset are conflicting, but in general are not very convincing (16–18). Given the interest in infectious agents as a trigger for PMR, our results suggesting higher PMR prevalence in rural areas might suggest a potential role for zoonotic pathogens (i.e., organisms that can be transmitted from wildlife or livestock to humans) in disease pathogenesis.

In many developed countries with aging populations, including Canada, decision makers are looking to administrative databases as a means of chronic disease surveillance to aid in resource planning. We were able to establish expected demographic patterns in our sample in terms of age and sex, as well as a novel finding of greater PMR prevalence in rural areas. This suggests that case ascertainment of some chronic rheumatic diseases using administrative data may indeed be feasible and useful. However, no method of case ascertainment, including those relying on administrative databases, is completely perfect. Our Bayesian hierarchical latent class regression model provides a means for addressing the imperfect nature of these data sources.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Bernatsky had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Bernatsky, Joseph, Pineau, Belisle, Clarke.
Acquisition of data. Bernatsky, Joseph, Lix.
Analysis and interpretation of data. Bernatsky, Joseph, Banerjee.

REFERENCES

1. Lawrence RC, Felson DT, Helmick CG, Arnold LM, Choi H, Deyo RA, et al, for the National Arthritis Data Workgroup. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States: part II. Arthritis Rheum 2008; 58:26–35.
2. Bartolome MJ, Martinez-Taboda VM, Lopez-Hoyos M, Blanco R, Rodriguez-Valverde V. Familial aggregation of polymyalgia rheumatica and giant cell arteritis: genetic and T cell repertoire analysis. Clin Exp Rheumatol 2001;19:259–64.
3. Doran MF, Crowson CS, O'Fallon WM, Hunder GG, Gabriel SE. Trends in the incidence of polymyalgia rheumatica over a 30 year period in Olmsted County, Minnesota, USA. J Rheumatol 2002;29:1694–7.
4. Salvarani C, Gabriel SE, O'Fallon WM, Hunder GG. Epidemiology of polymyalgia rheumatica in Olmsted County, Minnesota, 1970–1991. Arthritis Rheum 1995;38:369–73.
5. Bernatsky S, Joseph L, Pineau CA, Belisle P, Boivin JF, Banerjee D, et al. Estimating the prevalence of polymyositis and dermatomyositis from administrative data: age, sex, and regional differences. Ann Rheum Dis 2009;68:1192–6.
6. Joseph L, Gyorkos T, Coupal L. Bayesian estimation of disease prevalence and the prevalence of diagnostic tests in the absence of a gold standard. Am J Epidemiol 1995;141:263–72.
7. Dendukuri N, Joseph L. Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. Biometrics 2001;57:158–67.
8. Losina E, Barrett J, Baron JA, Katz JN. Accuracy of Medicare claims data for rheumatologic diagnoses in total hip replacement recipients. J Clin Epidemiol 2003;56:515–9.
9. Bernatsky S, Joseph L, Pineau CA, Tamblyn R, Feldman DE, Clarke AE. A population-based assessment of systemic lupus erythematosus incidence and prevalence: results and implications of using administrative data for epidemiological studies. Rheumatology (Oxford) 2007;46:1814–8.
10. Vacek PM. The effect of conditional dependence on the evaluation of diagnostic tests. Biometrics 1985;41:959–68.
11. Caporali R, Montecucco C, Epis O, Bobbio-Pallavicini F, Maio T, Cimmino MA. Presenting features of polymyalgia rheumatica (PMR) and rheumatoid arthritis with PMR-like onset: a prospective study. Ann Rheum Dis 2001;60:1021–4.
12. Gold LS, Ward MH, Dosemeci M, De Roos AJ. Systemic autoimmune disease mortality and occupational exposures. Arthritis Rheum 2007;56:3189–201.
13. Bovenzi M, Barbone F, Pisa FE, Betta A, Romeo L, Tonello A, et al. A case-control study of occupational exposures and systemic sclerosis. Int Arch Occup Environ Health 2004;77: 10–6.
14. Parks CG, Cooper GS, Nylander-French LA, Sanderson WT, Dement JM, Cohen PL, et al. Occupational exposure to crystalline silica and risk of systemic lupus erythematosus: a population-based, case–control study in the southeastern United States. Arthritis Rheum 2002;46:1840–50.
15. Olsson AR, Skogh T, Axelson O, Wingren G. Occupations and exposures in the work environment as determinants for rheumatoid arthritis. Occup Environ Med 2004;61:233–8.
16. Duhaut P, Bosshard S, Dumontet C. Giant cell arteritis and polymyalgia rheumatica: role of viral infections. Clin Exp Rheumatol 2000;18(4 Suppl 20):S22–3.
17. Nuti R, Giordano N, Martini G, Amendola A, Geraci S, Goutzamani J, et al. Is polymyalgia rheumatica caused by infectious agents? [letter]. J Rheumatol 2005;32:200–1.
18. Hemauer A, Modrow S, Georgi J, Helmke K, Vaith P, Lang B, et al. There is no association between polymyalgia rheumatica and acute parvovirus B19 infection [letter]. Ann Rheum Dis 1999;58:657.