Observation Period Effects on Estimation of Systemic Lupus Erythematosus Incidence and Prevalence in Quebec

Ryan Ng, Sasha Bernatsky, and Elham Rahme

ABSTRACT. ***Objective.*** To determine how duration of observation affects estimation of incidence and prevalence of systemic lupus erythematosus (SLE).

Methods. SLE incidence and prevalence estimates from data periods as brief as 3 years (2001-2003) were compared to estimates from a 15-year period (1989-2003).

Results. The 15-year period incidence was 5.6/100,000 (95% CI 5.0–6.1) and the prevalence was 59.1/100,000 (95% CI 57.4–60.8). When a 3-year period was used, incidence was overestimated by 238.1% and prevalence underestimated by 66.0%.

Conclusion. SLE incidence and prevalence estimates vary considerably according to the observation period; more than 5 years of data is likely required. (First Release June 15 2013; J Rheumatol 2013;40:1334–6; doi:10.3899/jrheum.121215)

Key Indexing Terms:
SYSTEMIC LUPUS ERYTHEMATOSUS INCIDENCE PREVALENCE

Health administrative databases are increasingly used for surveillance of rheumatic disease incidence and prevalence[1,2]. To accurately detect rheumatic cases, validated case ascertainment algorithms have been developed[1,3,4,5,6,7,8]. One overlooked limitation of these algorithms is how the duration of the observation period affects case detection. With fewer years of data, differentiating between incident and prevalent cases becomes difficult; also, additional prevalent cases will be missed. While systemic lupus erythematosus (SLE) observation period effects have been observed in the UK General Practice Research Database (GPRD)[9], these effects have never been formally quantified. To quantify SLE observation period effects, we estimated SLE incidence and prevalence over varying durations using population-based data for all medical encounters in Quebec, Canada.

From the Division of Clinical Epidemiology, McGill University Health Centre (MUHC); Division of Rheumatology, MUHC, Montreal, Quebec, Canada.

Supported by the Canadian Institutes of Health Research (CIHR). R. Ng was supported by a CIHR Research Fellowship. Dr. Bernatsky was supported by the CIHR, the Fonds de la recherche en santé du Québec, and the Research Institute of the MUHC, and is a Canadian Arthritis Network Scholar. E. Rahme is a senior investigator supported by the Fonds de la recherche en santé du Québec. She has received grants and consulting fees from Pfizer Inc. and Jenssen Inc. unrelated to this study.

R. Ng, MSc, Research Institute of the MUHC; E. Rahme, PhD, Clinical Epidemiology, MUHC; S. Bernatsky, MD, PhD, Division of Rheumatology, MUHC.

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

Accepted for publication April 22, 2013.

MATERIALS AND METHODS

Quebec health administrative databases covering all the health-insured Quebec residents (about 7.5 million of all ages) from January 1, 1989, to December 31, 2003, were used. The Med-Echo hospitalization database contains up to 16 diagnosis codes per admission. The Régie de l'Assurance Maladie du Québec (RAMQ; Quebec health insurance board) physician claims database contains 1 diagnosis code per patient visit. All diagnoses follow the *International Classification of Disease*, 9th revision (ICD-9) system. The RAMQ beneficiaries registration database contains demographic data.

The following algorithm, previously validated by chart abstraction (98.2% sensitivity, 72.5% specificity), was used to detect SLE[1,10,11]: (1) one SLE-coded (ICD-9: 710.0) discharge in the hospitalization database; and/or (2) one SLE-coded rheumatologist claim from the physician database; and/or (3) two SLE-coded nonrheumatologist claims at least 8 weeks apart, but within 2 years.

The index claim was a case's first SLE record, which could change depending on the observation period. The observation period was the time period for which data were available for analysis. From the maximum 15-year observation period (1989–2003), 12 shorter periods were constructed with lengths between 14 (1990–2003) and 3 years (2001–2003). Successive shorter periods were created by removing the earliest year.

Statistical analysis. The 2001 SLE incidence (and 95% CI) was calculated by dividing all 2001 incident cases by Quebec's non-SLE population on July 1, 2001[12]. A 2001 incident case had an index claim in that year. With longer observation periods, incident cases identified prior to 2001 were removed from the incident set. Therefore, with shorter periods, there would be more misclassified (as incident) prevalent SLE cases.

SLE prevalence in 2001 (and 95% CI) was calculated by dividing all living cases identified prior to December 31, 2001, by Quebec's population on July 1, 2001[12]. With longer observation periods, cases identified prior to 2001 only were added to the prevalent case set. Therefore, calculating prevalence from shorter periods would result in more undetected SLE cases. Analyses were performed using SAS 9.2.

RESULTS

Over the 15-year period, 4425 SLE cases were identified, 418 of which were in 2001. The average SLE diagnosis age

was 40.2 years (SD 17.5, range 0–89), and 84.4% of the patients were women. The 2001 incidence and prevalence were 5.6 (95% CI 5.0–6.1) and 59.1 (95% CI 57.4–60.8) SLE cases per 100,000 residents, respectively.

Incidence increased with shorter periods (Figure 1, dotted line). Compared to the 15-year period, the 2001 incidence was similar when the 14-year period was used, but increased slightly (6.1/100,000, 95% CI 5.5–6.6) with the 10-year period. With a 5-year period, 44% of prevalent cases were misclassified as incident (8.0/100,000, 95% CI 7.4–8.7), and even more with the 3-year period (20.1/100,000, 95% CI 19.1–21.1). Incidence estimates from periods shorter than 7 years did not overlap with the 15-year period estimate (Figure 1, hatched bars).

SLE prevalence estimates decreased with shorter periods (Figure 1, solid line). Compared to the 15-year period, prevalence decreased minimally with the 14-year period (58.4/100,000, 95% CI 56.7–60.2) and more (54.6/100,000, 95% CI 53.0–56.3) with the 10-year period. With the 5-year period, 39% of cases were undetected (46.0/100,000, 95% CI 34.7–37.4), and with the 3-year period, 66.0% (20.1/100,000, 95% CI 19.1–21.1) were undetected. All prevalence estimates from periods shorter than 11 years had no overlap with the 15-year period estimate (Figure 1, solid bars).

DISCUSSION

The 15-year SLE incidence and prevalence estimates are comparable, but slightly lower, versus other published incidence (2–8/100,000) and prevalence (20–240/100,000) estimates[13,14]. Some of this variation may be attributed to demographic or study method differences. While race/ethnicity is not discernible from our data, Quebec does not have large black or First Nations populations, 2 ethnicities with increased risk of developing SLE[4,15,16,17,18].

Our results show incidence and prevalence estimate variations related to the observation period. Based on our findings, to optimize case detection, the observation period should be maximized. A longer period may capture additional mild and/or inactive chronic disease cases (or those with less access to care) because of the increased likelihood of disease manifestation and associated health services use in these groups. For a 5-year period tradeoff, a 10-year period only misestimated incidence and prevalence by 8%. Using periods of 5 years or shorter should be avoided. However, if a short observation period is necessary because of limited data or secular trends, reasonable incidence and prevalence estimates might still be obtained with an adjustment factor that accounts for patients with no captured SLE-related health services use. For example, our 5-year prevalence estimate of 36/100,000 missed 39% of all patients with SLE (based on our 15-year prevalence estimate), so a 64% adjustment factor would allow researchers to still obtain a reasonable prevalence estimate.

One study limitation is that the data spanned only 15 years, so missing and/or misclassified SLE cases are still possible. Being the first study to our knowledge to quantify SLE observation period effects, the generalizability of our

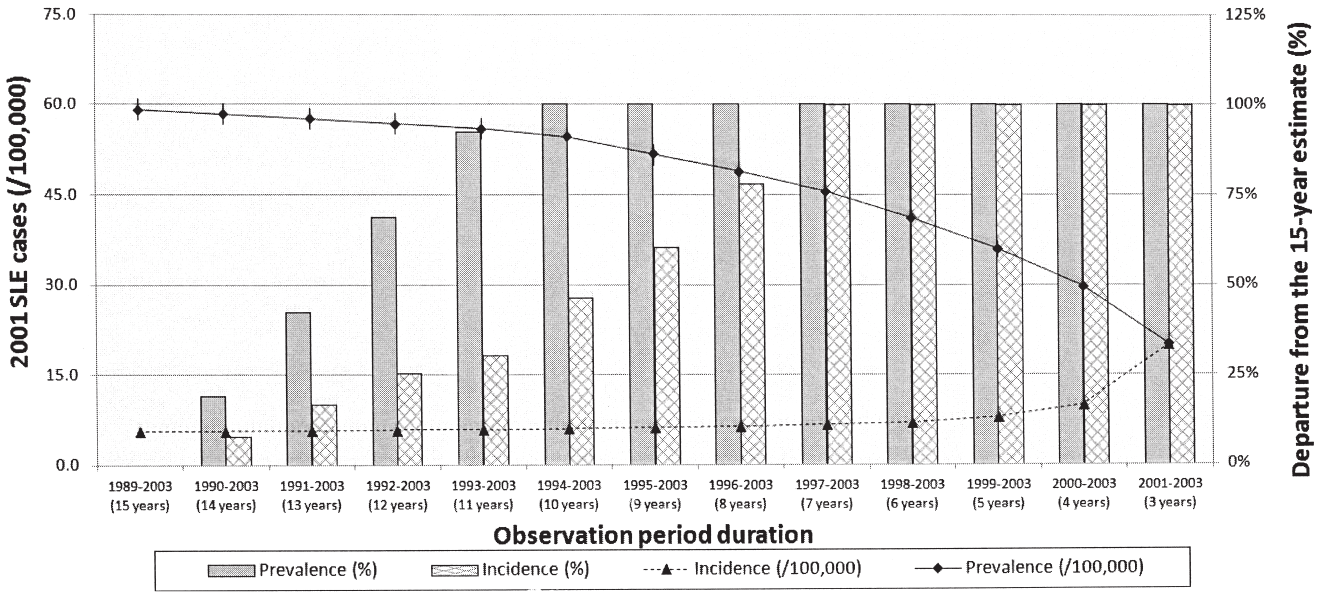

Figure 1. Systemic lupus erythematosus (SLE) 2001 incidence (dotted line) and prevalence (solid line) across different observation periods. As the observation period shrinks, there are fewer years to look for prior cases, so the incidence and prevalence estimates approach each other. The incidence and prevalence estimates become equivalent when there are no years remaining to look for prior cases. The solid gray bars show the departure of the shorter observation period's 2001 prevalence estimate from the 15-year observation period's prevalence estimate. The hatched bars show the departure of the shorter observation period's 2001 incidence estimate from the 15-year observation period's incidence 95% CI. The departures from the 15-year estimate were calculated by taking the percentage of the shorter observation period's 95% CI not bounded in the 15-year observation period's 95% CI.

results to other health administrative databases is relatively unknown. One comparison can be made to Nightingale, *et al*, who found at least a 10/100,000 SLE prevalence difference among women registered in the GPRD for 3 to 5 years versus 6 more years[9]. However, our data observed only a 5.2/100,000 prevalence difference between a 5- and 6-year period among all patients. We suspect that observation period effects exist for other diseases in other health administrative databases, but this requires confirmation.

Our study illustrates how estimates of incidence and prevalence of SLE are observation period–dependent, with a 10-year period being sufficient. Similar period requirements may exist for other chronic, relapsing-remitting diseases such as rheumatoid arthritis and multiple sclerosis, but additional research is needed. Regardless, careful consideration of the observation period is needed for all health administrative database research.

REFERENCES

1. 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 2007;46:1814-8.
2. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol 2005;58:323-37.
3. Public Health Agency of Canada, Arthritis Consumer Experts, Arthritis Community Research & Evaluation Unit, Canadian Arthritis Patients Alliance, Canadian Arthritis Network, Canadian Institute for Health Information, et al. Life with arthritis in Canada: A personal and public health challenge. 2010. [Internet. Accessed May 7, 2013.] Available from: www.phac-aspc.gc.ca/cd-mc/arthritis-arthrite/lwaic-vaaac-10/pdf/arthritis-2010-eng.pdf
4. Peschken CA, Esdaile JM. Systemic lupus erythematosus in North American Indians: A population based study. J Rheumatol 2000;27:1884-91.
5. Siegel M, Holley HL, Lee SL. Epidemiologic studies on systemic lupus erythematosus. Comparative data for New York City and Jefferson County, Alabama, 1956-1965. Arthritis Rheum 1970;13:802-11.
6. Amor B, Bouchet H, Delrieu F. [National survey on reactive arthritis by the French Society of Rheumatology]. Rev Rhum Mal Osteoartic 1983;50:733-43.
7. Chakravarty EF, Bush TM, Manzi S, Clarke AE, Ward MM. Prevalence of adult systemic lupus erythematosus in California and Pennsylvania in 2000: Estimates obtained using hospitalization data. Arthritis Rheum 2007;56:2092-4.
8. Michet CJ Jr, McKenna CH, Elveback LR, Kaslow RA, Kurland LT. Epidemiology of systemic lupus erythematosus and other connective tissue diseases in Rochester, Minnesota, 1950 through 1979. Mayo Clin Proc 1985;60:105-13.
9. Nightingale AL, Farmer RD, de Vries CS. Systemic lupus erythematosus prevalence in the UK: Methodological issues when using the General Practice Research Database to estimate frequency of chronic relapsing-remitting disease. Pharmacoepidemiol Drug Saf 2007;16:144-51.
10. Bernatsky S, Linehan T, Hanly JG. The accuracy of administrative data diagnoses of systemic autoimmune rheumatic diseases. J Rheumatol 2011;38:1612-6.
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 2011;31:549-54.
12. Statistics Canada. Estimates of total population for census divisions (CDs): Quebec, 2001-2004 (2004 Annual Demographic Statistics) [machine readable data file]. Ottawa, ON: Statistics Canada, Depository Services Program; 2005.
13. Lim SS, Drenkard C. Epidemiology of systemic lupus erythematosus: Capturing the butterfly. Curr Rheumatol Rep 2008;10:265-72.
14. Furst DE, Clarke AE, Fernandes AW, Bancroft T, Greth W, Iorga SR. Incidence and prevalence of adult systemic lupus erythematosus in a large US managed-care population. Lupus 2013;22:99-105.
15. Morton RO, Gershwin ME, Brady C, Steinberg AD. The incidence of systemic lupus erythematosus in North American Indians. J Rheumatol 1976;3:186-90.
16. Johnson AE, Gordon C, Palmer RG, Bacon PA. The prevalence and incidence of systemic lupus erythematosus in Birmingham, England. Relationship to ethnicity and country of birth. Arthritis Rheum 1995;38:551-8.
17. McCarty DJ, Manzi S, Medsger TA Jr, Ramsey-Goldman R, LaPorte RE, Kwoh CK. Incidence of systemic lupus erythematosus. Race and gender differences. Arthritis Rheum 1995;38:1260-70.
18. Hochberg MC. The incidence of systemic lupus erythematosus in Baltimore, Maryland, 1970-1977. Arthritis Rheum 1985;28:80-6.