Representativeness of patients and providers in the Canadian Primary Care Sentinel Surveillance Network: a cross-sectional study

John A. Queenan PhD, Tyler Williamson PhD, Shahriar Khan MSc, Neil Drummond PhD, Stephanie Garies MPH, Rachael Morkem MSc, Richard Birtwhistle MD

Abstract

Background: The Canadian Primary Care Sentinel Surveillance Network (CPCSSN) has established a national repository of primary care patient health data that is used for both surveillance and research. Our main objective was to determine how representative the data for patients and primary care practitioners in the CPCSSN are when compared with the Canadian population.

Methods: In this cross-sectional study, we compared the 2013 CPCSSN patient sample with age and sex information from the 2011 census. The CPCSSN provider sample in 2013 was compared with the 2013 National Physician Survey. Results were stratified by 5 clinically relevant age categories and sex, and male-to-female ratios were calculated.

Results: Patients who were 65 years of age and older represented 20.4% of the CPCSSN sample but only represented 14.8% of the Canadian population (2011 census). Among young adults (20–39 yr), 39.3% fewer men than women visited their primary care practitioner within 2 years. CPCSSN sample practitioners were more likely to be under 45 years of age, more likely to be female and more likely to be in an academic practice.

Interpretation: It is important to consider adjusting for age and sex when using CPCSSN data. CPCSSN practitioners are likely not representative of family physicians; therefore, CPCSSN needs to recruit more nonacademic practices, community clinics and practices that have a larger proportion of male providers.

The clinical data contained within electronic medical records (EMRs) are a valuable source of information for surveillance, research and practice quality improvement.[1,2] Issues of generalizability are a primary concern for observational studies that rely on data obtained from databases containing population-based patient health information.[3] A recent systematic review of the literature published from data held by the General Practice Research Database in the United Kingdom underscored the importance of generalizability and its subsequent impact on the validity of reported results.[4] The Canadian Primary Care Sentinel Surveillance Network (CPCSSN) is a network of 11 practice-based primary care research networks in British Columbia, Alberta, Manitoba, Ontario, Quebec, Nova Scotia, Newfoundland and Labrador, and the Northwest Territories. The CPCSSN established a process to extract, clean and merge EMR data from 12 different EMR databases into a single structured database of useable, quality primary care data.[5] Family physicians and nurse practitioners who contribute de-identified patient data to CPCSSN participate on a voluntary basis.

The scope of the data collected is comprehensive and robust, and has the potential to become a large, useable data source for Canadian health research, similar to well-established general practice databases in other countries.[6] For research and surveillance findings to be generalizable in a provincial or national context, it is important to consider how representative the data are to the Canadian base populations of primary care patients and practitioners.

The main objective of this study was to determine how representative the data of the patients and primary care practitioners in the CPCSSN are when compared with the Canadian population.

Methods

Setting

CPCSSN received approval for this study from the Research Ethics Board of each network's host university and from the Health Canada Research Ethics Board. Written consent for the collection and analysis of anonymous EMR data was obtained from all participating CPCSSN primary care providers.

The first wave of recruitment took place between 2008 and 2010 and focused on recruiting family practices that were associated with academic or university sites, where providers were more amenable to participating in research. As CPCSSN became more established, recruitment strategies were broadened to include nonacademic practices in various settings (urban, suburban and rural). The resulting sample of patients, providers and practices is both a convenience and purposive sample. Patients are given the opportunity to opt out of the CPCSSN database except in Quebec, where provincial law mandates an opt-in process. At present, only a small number of patients requested exclusion (*n* = 1) or refused to participate (*n* = 314). Our database has grown since its inception in 2008 to include more than 800 practitioners across Canada who contribute data on over a million patients.

CPCSSN data are from all patient encounters that are recorded in the EMR. We extract most data elements for all patients, including demographics, diagnoses, laboratory and examination results, medications, medical procedures, risk factors, referrals and physician billing. CPCSSN has validated methods to estimate the prevalence of 8 chronic conditions.[7]

Sources of data and study population

The study design was cross-sectional. We compared select information about the CPCSSN patient and provider samples with publicly available information from the 2011 census[8] and the 2013 National Physician Survey.[9] The CPCSSN patient sample was defined as all patients in the database as of Sept. 30, 2013, who had had at least 1 clinic encounter in the previous 2 years. No other exclusion or inclusion criteria were applied. This method was used previously as a reasonable estimate of the "active practice."[10] The census is the most comprehensive source of Canadian demographic data and is readily available.

The National Physician Survey is sent annually to all licensed Canadian physicians and provides information on both primary care and specialist providers. Retired physicians, medical students and residents are not included. The CPCSSN primary care provider population was compared with the primary care respondents to the 2013 National Physician Survey. Although the National Physician Survey has low response rates, it was the only suitable comparator available in Canada.

Patient age and sex comparisons were also performed at the provincial level. Information on CPCSSN practitioners' age, sex, urban or rural clinic location and if they practised in an academic setting were provided by the network sites and compared with the 2013 National Physician Survey.

Statistical analysis

Eligible patients in the CPCSSN sample (2-yr contact) were stratified by clinically relevant age groups (infants 0–4 yr,

children 5–19 yr, young adults 20–39 yr, middle-aged adults 40–64 yr and older adults 65 yr and older) and sex. The proportion of patients in each stratum, 95% confidence intervals (CIs) and the ratio of males to females were calculated. The comparison age and sex distributions of the Canadian population were similarly stratified. CPCSSN practitioners were stratified according to age (< 35, 35–44, 45–54, 55–64 and ≥ 65 yr), sex, practice setting (academic or community practice) and rural or urban practice location (rurality was indicated by 0 in the second digit of the postal code). These strata were chosen to correspond to those that could be obtained from the 2013 National Physician Survey. Descriptive data analysis was performed using SAS version 9.3.

Results

Table 1 shows the demographic characteristics of the patients in the CPCSSN database. As of Sept. 30, 2013, 392 443 patients who had at least 1 encounter with their primary care provider in the last 2 years and 478 providers (465 family physicians, 14 nurse practitioners and 9 with provider type data missing) were registered with CPCSSN. The patients in the CPCSSN were, on average, 43 years of age (SD 24 yr), and there were more women than men (57.1% v. 42.9%). About

Table 1: Characteristics of patients in the CPCSSN database with a clinical encounter in the previous 2 years (*n* = 392 443; as of Sept. 30, 2013) compared with the 2011 Canadian Census population (full census data, including non-CPCSSN provinces, *n* = 33 476 715)

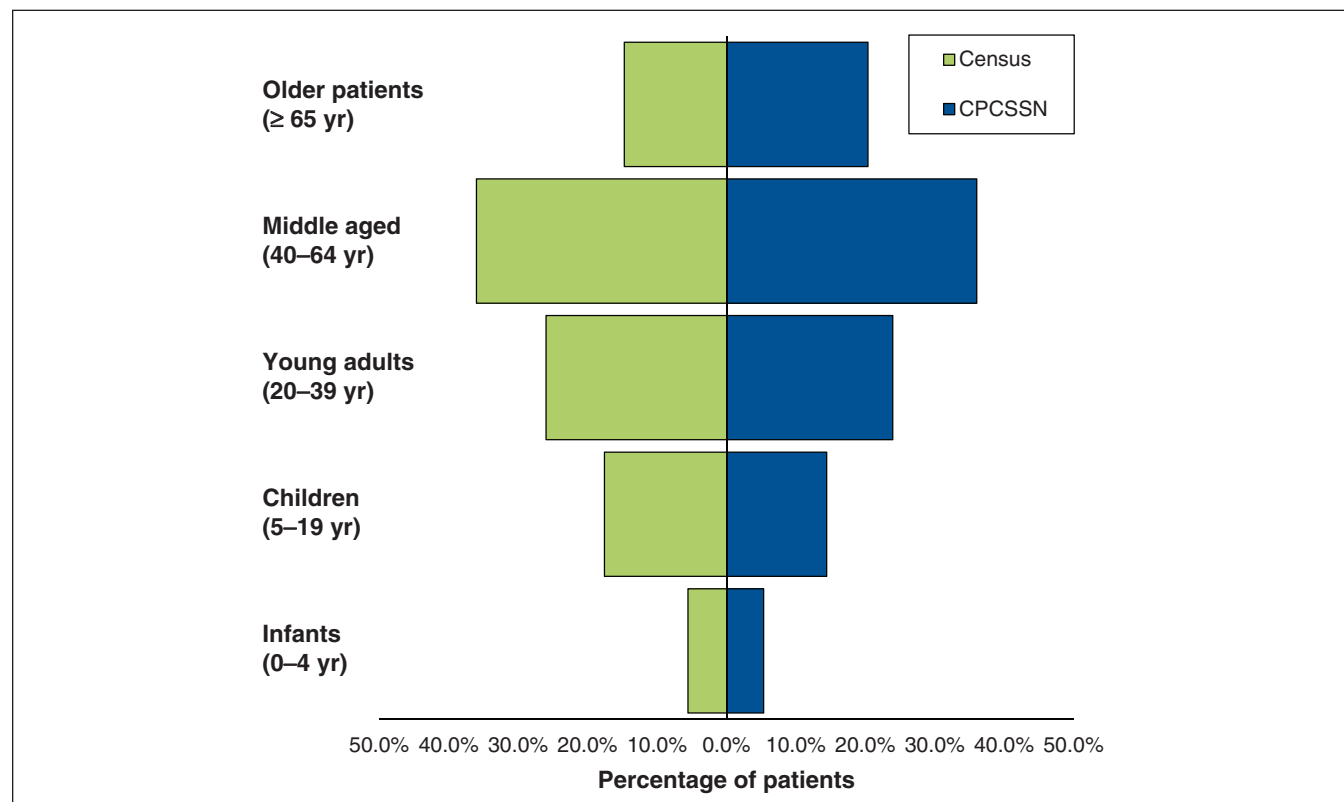| Characteristic	CPCSSN	Census 2011[8]
Age, yr		
Mean ± SD	43.1 ± 23.5	39.6 ± 22.6
Median (IQR)	44.0 (24.0–61.0)	40.0 (21.0–56.0)
Sex, %		
Male	42.9	49.0
Female	57.1	51.0
Residence location, %		
Urban	77.1	81.0
Rural	22.9	19.0
Province,* %		
British Columbia	2.6	13.1
Alberta	15.8	10.9
Manitoba	8.2	3.6
Ontario	48.2	38.4
Quebec	1.6	23.6
Nova Scotia	14.1	2.8
Newfoundland	9.5	1.5

Note: CPCSSN = Canadian Primary Care Sentinel Surveillance Network, IQR = interquartile range.
*Census total does not add up to 100.0% because non-CPCSSN provinces were excluded.

one-quarter (22.9%) of the patients had a rural postal code. For the CPCSSN sample, Ontario contributed 48.2%, with Alberta, Nova Scotia, and Newfoundland and Labrador making the next largest contributions (15.8%, 14.1% and 9.5%, respectively). Networks in British Columbia and Quebec made very modest contributions to the overall sample (2.6% and 1.6%, respectively), since the growth of CPCSSN in those provinces has been slow because BC is a relatively new network and health legislative requirements in Quebec slow the process. Territorial data were not available as of Sept. 30, 2013.

Figure 1 presents the population pyramid comparing the 2013 CPCSSN sample and the 2011 Census. The CPCSSN patient population was, on average, older than the baseline Canadian population. The largest differences were for the

Figure 1: Population pyramid comparing the 2013 CPCSSN sample population and the 2011 Canadian census, grouped by clinically relevant age categories. CPCSSN = Canadian Primary Care Sentinel Surveillance Network.

Table 2: Sex-stratified proportions (95% CIs) of the Canadian population (2011 full census data) compared with the national sample patient population in the CPCSSN

Characteristic by age group	Census 2011 population N = 33 476 715			CPCSSN national population N = 392 443		
	No. of males, % (95% CI) n = 16 414 225	No. of females, % (95% CI) n = 17 062 490	Sex ratio	No. of males, % (95% CI) n = 168 316	No. of females, % (95% CI) n = 224 127	Sex ratio
Infants, 0–4 yr	961 150, 5.85 (5.84–5.87)	915 950, 5.37 (5.36–5.38)	1.05	10 492, 6.23 (6.12–6.35)	10 390, 4.64 (4.55–4.72)	1.01
Children, 5–19 yr	3 025 790, 18.43 (18.41–18.45)	2 882 605, 16.89 (16.88–16.91)	1.05	27 582, 16.39 (16.21–16.56)	28 886, 12.89 (12.75–13.03)	0.95
Young adults, 20–39 yr	4 309 075, 26.25 (26.23–26.27)	4 384 810, 25.70 (25.68–25.72)	0.98	35 434, 21.10 (20.86–21.25)	58 415, 26.06 (25.88–26.25)	0.61
Middle aged, 40–64 yr	5 919 450, 36.06 (36.04–36.09)	6 132 845, 35.94 (35.92–35.97)	0.97	60 236, 35.79 (35.56–36.02)	81 128, 36.20 (36.00–36.40)	0.74
Older adults, 65 yr and older	2 198 760, 13.39 (13.38–13.41)	2 746 280, 16.09 (16.08–16.11)	0.80	34 572, 20.54 (20.35–20.73)	45 308, 20.21 (20.05–20.38)	0.76

Note: CI = confidence interval, CPCSSN = Canadian Primary Care Sentinel Surveillance Network, sex ratio = no. of males/no. of females.

65 years and older and the less than 30 years of age groups. Patients 65 years and older made up 20.4% of the CPCSSN patient population, but they made up only 14.8% of the Canadian population according to the census. Conversely, the CPCSSN patient sample slightly underrepresented young adults and children compared with the Canadian population (23.9% v. 26.0% and 14.4% v. 17.6%, respectively).

Table 2 compares the 2011 census stratified by sex. The ratio of males to females was close to parity in both the infant and children age categories. The largest difference was observed in the young adult population, in which 39.3% fewer men visited their physician within 2 years compared with women. There were significantly fewer men than women in the middle age and older adult categories (25.8% and 23.7% fewer, respectively).

Table 3 provides comparisons of the demographic information for CPCSSN practitioners and respondents to the 2013 National Physician Survey. CPCSSN practitioners were younger, more likely to be female (51.1% v. 44.0%) and from academic practices (19.3% v. 7.8%) compared with the National Physician Survey respondents.

Geographic location of participating practices was similar (15.9% rural clinics v. 16.0%) (data not shown).

Additional population pyramids are available in Appendix 1 (available at www.cmajopen.ca/content/4/1/E28/suppl/DC1). Figure S1 shows the population pyramid comparing the 2013 CPCSSN sample and the 2011 Canadian census, and Figure S2 shows the national population pyramids for men and women. Figures S3–9 show the analogous population pyramids for each participating province. All of the provinces, with the exception of Quebec, had similar distributions when compared with their national and provincial data (Figure S1). Quebec had the least number of patients contributing data to the CPCSSN database (men n = 2501, women n = 3871).

Interpretation

CPCSSN patients are only somewhat representative of the Canadian general population. Provincial-level comparison was reasonable, except for Quebec and British Columbia (contributions of 1.6% and 2.6%, respectively). This study highlights areas for improvement and the need to have practice participation from all of Canada to improve overall representativeness. Past primary care research suggests that Canadians who visit their family physician are older and more likely to be women.[11–13] This pattern was repeated in the CPCSSN patient population and is typical of most primary care study groups.[11–13] The underlying reason for the observed differences may be that women seek medical care for themselves or for their children more frequently than men.[11–13] Therefore, it is important to consider adjustments in age and sex for any analysis conducted using CPCSSN data. Although the Quebec network had a lower number of males aged 20–39 years compared with the other provincial networks, it is possible that this difference was due to the smaller sample size that Quebec contributed (1.6%). The representation of CPCSSN at the provincial level will improve with increasing numbers of both patients and practi-

tioners from across the country. As the numbers of CPCSSN practitioners and patients continue to grow by strategic sampling, issues surrounding representativeness will decrease.

CPCSSN practitioners may not be representative of family physicians who responded to the 2013 National Physician Survey. In general, CPCSSN practitioners are younger and practice more often in an academic centre than their colleagues in family medicine. As interest increases in understanding the contributions made by nonacademic physicians and other health professionals contributing to team-based primary health care, representativeness among practices at the general-function or even specific-task level becomes increasingly relevant and desirable. Therefore, CPCSSN will continue to recruit with an emphasis on nonacademic physicians to ensure future representativeness.

Limitations

There were no similar studies that we could find that reported on the external validity/generalizability of a national EMR. The literature has focused on issues of internal validity and its subsequent impact on external validity.

Because CPCSSN practitioners, and by association their patients, were not randomly selected, the extent to which

Table 3: Representativeness of physicians in the national CPCSSN data (up to Sept. 30, 2013) compared with respondents to the 2013 NPS

Characteristic	No. (%) of CPCSSN physicians n = 478	No. (%) of 2013 NPS respondents n = 31 799
Age, yr		
< 35	59 (13.7)	3 434 (10.8)
35–44	110 (25.5)	6 805 (21.4)
45–54	125 (29.0)	8 554 (26.9)
55–64	104 (24.1)	9 000 (28.3)
≥ 65	33 (7.7)	4 007 (12.6)
No. missing	47	0
Sex		
Male	234 (48.9)	17 807 (56.0)
Female	244 (51.0)	13 992 (44.0)
Practice location		
Urban/nonrural	402 (84.1)	26 432 (84.0)
Rural	76 (15.9)	5 035 (16.0)
No. missing	0	332
Practice setting		
Academic primary care	92 (19.3)	1 982 (7.8)
Nonacademic primary care	386 (80.7)	23 423 (92.2)
No. missing	0	6 394

Note: CPCSSN = Canadian Primary Care Sentinel Surveillance Network, NPS = National Physician Survey.

results from within CPCSSN can be generalized to Canadian primary care is related to the extent to which CPCSSN practitioners and patients reflect their respective groups. Currently, we are unable to assess the magnitude of bias in our sample, and future methodological work may be able to quantify to what extent selection bias is present. Although we can have reasonable assurance that age- and sex-adjusted prevalence rates may be generalizable to the Canadian primary care patient population, we have less confidence in our ability to calculate prevalence rates within socioeconomic strata or by ethnicity, since these variables are not well recorded in EMRs. In addition, because the data held within the CPCSSN database are anonymized, we were unable to assess to what extent an individual patient may be present in 1 or more physician rosters. We are also cautious about reporting age-adjusted prevalence rates for individual provinces, especially for BC and Quebec, which contributed the least number of patients.

In using the National Physician Survey to compare our physician population with that of the rest of Canada, the low response rate of the National Physician Survey may have underestimated the congruence between CPCSSN sentinels and the actual population of primary care providers. The national rate of the National Physician Survey was 17.0%, and the provincial rates ranged from 12.8% to 22.9%.[9] Therefore, we are unable to confidently assess to what extent the physicians in the CPCSSN represent the population of family physicians in Canada.

Despite these important limitations, CPCSSN currently provides comprehensive information on over 1 million patients.

Conclusion

It is important to consider adjusting for age and sex when using CPCSSN data. Representativeness of CPCSSN data will continue to be of paramount consideration. In Canada, provincial contexts are of central importance to health care system planning, organization and delivery. This study revealed priority areas for improvement in the CPCSSN: we need to target non-academic practices and community clinics more often than academic ones and target practices that have a larger proportion of male providers. However, some aspects will continue to be difficult to reconcile (e.g., younger men are less likely to see a family care practitioner and older practitioners are more likely to use paper charts). CPCSSN's ability to undertake analyses of data that are representative at the national, provincial or territorial level, as well as primary care practitioners and patients at those levels, is crucial to its sustainability. The solutions to improved representativeness will lie in the further expansion of the database in all locations to effectively accommodate multiple levels. Adopting the same purposive sampling strategies used in our past recruitment in relation to national representativeness and incorporating them into our future local-recruitment strategies will help to ensure representativeness at the provincial and regional levels among providers and patients.

References

1. Garg AX, Adhikari NKJ, McDonald H, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA* 2005;293:1223-38.
2. Ludwick DA, Doucette J. Adopting electronic medical records in primary care: Lessons learned from health information systems implementation experience in seven countries. *Int J Med Inform* 2009;78:22-31.
3. Lawrenson R, Williams T, Farmer R. Clinical information for research; the use of general practice databases. *J Public Health Med* 1999;21:299-304.
4. Herrett E, Thomas SL, Schoonen WM, et al. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol* 2010;69:4-14.
5. Birtwhistle R, Keshavjee K, Lambert-Lanning A, et al. Building a pan-Canadian primary care sentinel surveillance network: initial development and moving forward. *J Am Board Fam Med* 2009;22:412-22.
6. Lawson DH, Sherman V, Hollowell J. The General Practice Research Database. Scientific and Ethical Advisory Group. *QJM* 1998;91:445-52.
7. Williamson T, Green ME, Birtwhistle R, et al. Validating the 8 CPCSSN case definitions for chronic disease surveillance in a primary care database of electronic health records. *Ann Fam Med* 2014;12:367-72.
8. *2011 Census profile*. Ottawa: Statistics Canada; 2011. Available: www12.statcan.gc.ca/census-recensement/2011/dp-pd/prof/index.cfm?Lang=E (accessed 2016 Jan. 14).
9. *2013 survey results*. Mississauga (ON): National Physician Survey 2013; 2013. Available: http://nationalphysiciansurvey.ca/surveys/2013-survey/survey-results (accessed 2015 Oct. 15).
10. Greiver M, Williamson T, Barber D, et al. Prevalence and epidemiology of diabetes in Canadian primary care practices: a report from the canadian primary care sentinel surveillance network. *Can J Diabetes* 2014;38:179-85.
11. Nie JX, Wang L, Tracy CS, et al. Health care service utilization among the elderly: findings from the Study to Understand the Chronic Condition Experience of the Elderly and the Disabled (SUCCEED project). *J Eval Clin Pract* 2008; 14:1044-9.
12. Bertakis KD, Azari R, Helms L, et al. Gender differences in the utilization of health care services. *J Fam Pract* 2000;49:147-52.
13. Mustard CA, Kaufert P, Kozyrskyj A, et al. Sex differences in the use of health care services. *N Engl J Med* 1998;338:1678-83.

Affiliations: Department of Family Medicine (Queenan, Khan, Drummond, Morkem, Birtwhistle), Queen's University, Kingston, Ont.; Community Health Sciences (Williamson, Garies), University of Calgary, Calgary, Alta.