

International Journal of Population Data Science

Journal Website: www.ijpds.org



Trend control charts for multiple sclerosis case definitions

Naomi C. Hamm^{1,*}, Ruth Ann Marrie^{1,2}, Depeng Jiang¹, Pourang Irani³, and Lisa M. Lix¹

Submission History

Submitted:	13/09/2023
Accepted:	08/02/2024
Published:	30/04/2024

¹Department of Community Health Sciences, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada

²Department of Internal Medicine, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada

³Department of Computer Science, Mathematics, Physics and Statistics, Irving K. Barber Faculty of Science, University of British Columbia, Kelowna, British Columbia, Canada

Abstract

Introduction

The validity of chronic disease case definitions for administrative health data may change over time due to changes in data quality. Trend control charts to identify out-of-control (OOC; i.e., unexpected) observations in a time series may indicate where disease estimates are influenced by changes in data quality.

Objective

Apply and compare trend control charts methods for multiple sclerosis (MS) incidence and prevalence estimates using previously-validated case definitions for Manitoba, Canada.

Methods

Eight case definitions were identified from published literature and applied to Manitoba administrative health data from January 1, 1972 to December 31, 2018. Incidence and prevalence trends were modeled using negative binomial and generalized estimating equation models, respectively. Trend control charts were used to plot predicted case counts against observed case counts. Control limits to identify OOC observations were calculated using two methods: predicted case count $\pm 0.8 \times \text{standard deviation}$ ($0.8 \times \text{SD}$) and predicted case count $\pm 2 \times \text{standard deviation}$ ($2 \times \text{SD}$). Differences in proportion of OOC observations across case definitions was assessed using McNemar's test.

Results

The proportion of OOC observations ranged from 0.71 to 0.90 for incidence and 0.72 to 0.98 for prevalence when using the $0.8 \times \text{SD}$ control limits. A lower proportion of OOC observations (0.46 to 0.74 for incidence; 0.30 to 0.74 for prevalence) was observed for the $2 \times \text{SD}$ control limits. Neither method resulted in significant differences in OOC observations across case definitions.

Conclusions

The proportion of OOC observations in trend control charts varied with the control limit method adopted, but statistical significance did not. Trend control charts are a potentially useful tool for developing surveillance methods, but may benefit from disease-specific calibrated control limits.

Keywords

control charts; International Classification of Diseases; multiple sclerosis; prevalence trends; incidence trends

*Corresponding Author:

Email Address: lettn@myumanitoba.ca (Naomi C. Hamm)



Introduction

Administrative health data are widely used for chronic disease research and surveillance due to their population and temporal coverage, objective measurement of health conditions, and low cost to access. However, these data were not originally collected for research or surveillance; there are multiple challenges associated with using administrative health data for chronic disease research and surveillance. Changes in the coding systems used to ascertain health conditions (i.e. changes to International Classification of Disease [ICD] codes) [1, 2], as well as the processes used to generate those codes (i.e., clinical guidelines, coding practices, diagnostic criteria, and healthcare processes) [3], may be problematic when estimating incidence and prevalence trends. For example, ICD codes are updated periodically by the World Health Organization to reflect changes in the medical field; codes from different versions do not always map directly to one another [4]. Compared to the ninth revision (i.e., ICD-9), the tenth revision (i.e., ICD-10) has almost twice the number of codes and some conditions have been grouped into different chapters [5]. These changes in the ICD coding system can impact population health measurement over time [1–3].

Case definitions, which are rules or algorithms based on ICD codes, are used to obtain disease estimates from population-based administrative health data [6, 7]. Constructing case definitions involves careful consideration of the disease criteria (i.e., disease treatment duration and frequency, screening and diagnostic procedures, and time between disease onset and diagnosis) and how they map onto the available data elements (i.e., ICD codes) [8, 9]. Often, multiple case definitions are built and validated within the same study. Validation involves comparing case definition results to information from external data sources. This is usually done at a single point in time and rarely re-visited over time. It may not be reasonable to assume that validity measures, such as sensitivity or specificity, remain constant over time [10]. In addition, measures that are influenced by disease prevalence, such as positive predictive value and negative predictive value, are likely to change as population health changes. This, along with changes in data quality for disease identification due to changes in coding systems [4] or other data generation processes, may influence case definition performance. Therefore, changes in disease trend estimates produced using validated case definitions may reflect changes in data quality for disease case identification rather than true changes in population health [11, 12].

Control charts are a statistical approach to monitor processes over time and distinguish between random and non-random sources of error [13]. Control limits are set around a centre (i.e., average) line, and observations outside the control limits are deemed 'out-of-control' (OOC). These OOC observations suggest non-random sources of variation have influenced the process of interest [13]. By flagging (i.e., identifying) OOC observations, control charts can indicate where non-random influences, such as a change in policy/procedure, intervention, or unexpected factors (e.g., COVID-19 pandemic), may impacted the process of interest. Choice of control limits to identify OOC observations may vary based on the process of interest [14] and can be determined based on a clinical or statistical criteria (e.g.,

percent increase/reduction of relative risk or log-likelihood ratio), or both types of criteria [14, 15]. There is no single method to calculate control limits, and the choice of methods can influence the observations identified as OOC observations [14].

Control charts have been applied in various health settings to monitor the quality of healthcare delivery, assess population health trends, and monitor public health surveillance programs [14, 16–19]. For disease trends obtained from administrative health data, potential contributors to non-random error include changes in case definition performance due to variations in data quality over time. Therefore, researchers and epidemiologists can use trend control charts as a tool to inform how and where changes in data quality may affect case definition validity. Previously, we applied trend control charts to administrative health data for juvenile diabetes to describe trend stability for several validated case definitions [12].

Multiple sclerosis (MS) is a complex disease to diagnose, often requiring multiple healthcare provider visits [20]. At onset, MS is an episodic disease in about 85% of affected individuals, characterized by periods of relapse and remission [21]. Over time, many individuals develop gradually worsening disease symptoms even in the absence of relapse [21]. MS diagnostic criteria have changed over time as new technologies, such as magnetic resonance imaging (MRI), have become available [21]. The control chart methodology and control limits previously applied to trend estimates from juvenile diabetes case definitions may not be the best choice for all diseases. Differences in the shape of the trend, case ascertainment method, and magnitude of prevalence and incidence estimates may affect this choice. Therefore, this study applied trend control chart methods using different control limits to MS incidence and prevalence trends for previously published, validated MS case definitions. OOC observations from two different control limit calculations were identified and compared over time across case definitions.

Methods

Study data

Administrative health data from January 1, 1972 to December 31, 2018 were obtained from the Manitoba Population Research Data Repository housed at the Manitoba Centre for Health Policy. Manitoba has a universal healthcare system; therefore, this data repository contains all healthcare contacts covered by provincial healthcare insurance for Manitoba's 1.3 million residents (estimated 99% of population) [22]. Hospital records, contained in the Hospital Discharge Abstracts database, are coded using ICDA-8 (from January 1, 1972 to March 31, 1979), ICD-9-Clinical Modifications (CM) (from April 1, 1979 to March 31, 2004), and ICD-10-Canadian version (CA) (April 1, 2004 onwards). Physician records, contained in the Medical claims/Medical services database are coded using ICDA-8 (from January 1, 1972 to March 31, 1979) and ICD-9-CM (April 1, 1979 onwards). Hospital and physician records were originally coded in the seventh revision of ICD codes from 1972 to 1974 and later converted to ICD-8 by the data provider. We had no information about the conversion methods used by the data provider and therefore

excluded these years from the analysis. However, these data were used in the lookback period to ascertain incident cases (see *Study Cohort*). Health insurance coverage dates, sex, and date of birth were obtained from the Manitoba Health Insurance Registry.

Case definitions

Previously-validated case definitions for MS were identified from published literature. PubMed, Google Scholar, and Embase were searched up to December 1, 2020 using the search string “(((Multiple Sclerosis) AND (administrative health data OR claims data)) AND (case definition)) AND (incidence)”. Case definitions were selected if they used hospital and/or physician records and if at least one measure of validity (e.g., sensitivity) was reported. A total of eight validated case definitions were identified [23–29].

Study cohort

A separate cohort was created for each case definition. For case definitions with an observation window (i.e., the time period in which a specific type and number of ICD codes must occur in the data for an individual to be considered a disease case), individuals required healthcare insurance coverage during the entire observation window to be retained in the cohort. For case definitions without an observation window, individuals required at least one day of health insurance coverage to be retained in the cohort. Individuals were classified as cases if they met the case definition criteria. A 5-year look back period was used to ascertain incident cases [24], where an individual had to have no MS ICD codes within the last 5 years to be considered incident. Incident and prevalent counts were aggregated by sex and age group (20 to 29 years; 30 to 39 years; 40 to 49 years; 50 to 59 years; 60 to 69 years; 70 to 79 years; 80+ years).

Statistical analysis

Statistical methods followed those previously reported [12]. Briefly, incident case counts were modelled using negative binomial regression; prevalent case counts were modelled using generalized estimating equations (GEEs) regression with a Poisson distribution and a first order autoregressive correlation structure. Age group (reference group: 20 to 29 years), sex, and year were covariates. The model offset was the natural logarithm of the cohort size. Where trends were determined to be non-linear, a restricted cubic spline [30] was used.

Incidence and prevalence trends for each case definition were graphed by year using observed-expected trend control charts, where the observed value was the annual estimated case count, and the expected value was the model-predicted count summed across age groups and sexes. Control limits were calculated using two different methods. The first method identified OOC observations when there were large differences between model-predicted and observed counts. The size of the difference was defined using Cohen's effect size [31], where a large effect size is observed when the difference between two means divided by their pooled standard deviation is 0.8 or greater. Accordingly, control limits were calculated as the summed model-predicted count $\pm 0.8 \times$ pooled standard

deviation of the model-predicted counts. The second method identified OOC observations based on traditional control chart methods where control limits are based on 95% or 99% confidence intervals. Specifically, control limits were calculated as the summed model-predicted count $\pm 2 \times$ pooled standard deviation of the model-predicted count (approximately equal to the 95% confidence interval). For both methods, annual case counts were classified as OOC if they fell outside the calculated control limits. To compare trend stability over time, case counts from the years 1975 to 2017 were used for prevalence and 1977 to 2017 for incidence. Data after 2017 were truncated as the case definition with a two-year observation window did not have case counts beyond 2017. For incidence, data before 1977 were used in the 5-year look back period.

McNemar's test [32] was used to test for differences in the frequency of OOC observations for case definitions. Differences were tested across all years, by ICD version, and by years surrounding transitions from one ICD version to another. If the number of observations was less than 10, McNemar's exact test was used. The case definition of three or more hospital or physician MS diagnosis codes (ICD-9-CM: 340; ICD-10-CA: G35) was selected as the reference case definition to compare to all other case definitions as this was the most frequently identified case definition used in previous research. A Holm-Bonferroni adjustment [33] was used to control the overall probability of a Type I error. OOC proportion, calculated as the number of OOC years for a case definition divided by the number of study years (43 for prevalence; 41 for incidence), was also compared across case definitions.

All data analyses were performed using R version 4.1.0. The MASS package [34] was used to fit the negative binomial models and the geepack package [35] was used to fit the GEE models.

Results

Details on the case definitions, including their ICD codes and regions they have been validated in, can be found in Table 1. Most identified validation studies applied case definitions in Canadian regions. Case definitions applied in Sweden, Germany, and Hungary were identified as well. Only 2 of the 8 identified case definitions used observation windows. Three case definitions used 3 or more hospital or physician codes; however, there were variations on the ICD codes used for case identification.

Trend control charts for incidence and prevalence using both control limit calculations can be found in the supplemental material (Supplementary Figures 1–4). Incidence did not consistently increase or decrease for the case definitions. Prevalence increased relatively consistently over time across all case definitions.

The frequency and proportion of OOC annual counts obtained when using the $0.8 \times$ SD control limits are reported in Table 2, along with results from McNemar's test. For incidence, the proportion of OOC observations ranged from 0.71 to 0.90, with a mean of 0.84 across all case definitions. For prevalence, the proportion of OOC observations ranged from 0.72 to 0.98 with a mean value of 0.83 across all case definitions. McNemar's test with a

Table 1: Description of identified validated multiple sclerosis case definitions

Case definition name	Description	ICD-codes	Country/region of validation study
3+ (A)	3 or more hospital or physician codes	ICD-9-CM: 340 ICD-10-CA: G35	Canadian provinces: British Columbia, Manitoba, Quebec, and Nova Scotia, Saskatchewan Sweden
3+ (B)	3 or more hospital or physician codes	ICD-9-CM: 340 ICD-10-CA: G35 Incidence date determined using: ICD-9-CM: 341.0, 341.9, 323, 377.3, 323.82 ICD-10-CA: G36, G36.0, G37.9, G36.9, H46, G37	Canadian province: Nova Scotia
3+ (C)	3 or more hospital or physician codes	ICD-9-CM: 340 ICD-10: G35 Note: at least one MS must be submitted by neurologist. MS code had to occur in at least 2 different calendar years (not consecutive)	Hungary
7+	7 or more hospital or physician codes	ICD-9-CM: 340 ICD-10-CA: G35 Incidence date determined using: ICD-9-CM: 377.3, 323.82, 323, 341.9, 341.0 ICD-10-CA: H46, G37, G36.9, G37.8, G36, G36.0	Canadian province: Nova Scotia
7+ or 3+ (coverage)	7 or more hospital or physician codes if 3+ years of coverage 3 or more hospital or physician codes if <3 years of coverage	ICD-9-CM: 340 ICD-10-CA: G35 Incidence date determined using: ICD-9-CM: 377.3, 323.82, 323, 341.9, 341.0 ICD-10-CA: H46, G37, G36.9, G37.8, G36, G36.0	Canadian province: Nova Scotia
7+ or 3+ (date)	7 or more hospital or physician codes up to and including Dec 31, 1997 3 or more hospital or physician codes after Dec 31, 1997	ICD-9-CM: 340 ICD-10-CA: G35 Incidence date determined using: ICD-9-CM: 377.3, 323.82, 323, 341.9, 340, 341.0 ICD-10-CA: H46, G37, G36.9, G37.8, G36, G35, G36.0	Canadian province: Manitoba
1+ hosp or 2+ phys	1 or more hospital or 2 or more physician codes in 1 year	ICD-9-CM: 340 ICD-10-CA: G35	Germany
1+ hosp or 5+ phys	1 or more hospital or 5 or more physician codes in 2 years	ICD-9-CM: 340 ICD-10-CA: G35	Canadian province: Saskatchewan

Holm-Bonferroni adjustment revealed no statically significant differences between the reference case definition and the remaining case definitions for both incident and prevalent case counts.

The proportion of OOC observations obtained when using 2*SD control limits and results from the McNemar's tests are reported in Table 3. Using these control limits, a lower

proportion of observations were flagged as OOC (0.46–0.74 with a mean of 0.58 for incidence; 0.30–0.74 with a mean of 0.55 for prevalence across all case definitions). Similar to the 0.8*SD control limits, no statically significant differences between the reference case definition and the remaining case definitions were found for both incident and prevalent case counts.

Discussion

MS case definitions applied to Manitoba's administrative health data produced variable incidence trends and increasing prevalence trends. The proportion of OOC observations detected via trend control charts for incidence and prevalence ranged from 0.78 to 0.90 and 0.72 to 0.98, respectively when using 0.8*SD control limits. Wider control limits (2*SD) produced a lower proportion of OOC observations ranging from 0.46 to 0.71 for incidence and 0.30 to 0.74 for prevalence. No significant differences in the frequency of OOC observations were seen across case definitions for either methods of control limit calculation.

Changes in the quality of administrative health data for chronic disease research and surveillance can impact case definition performance over time. Trend control charts provide a means of identifying where changes in disease trends are potentially reflective of changes in data quality, rather than changes in population health. By applying trend control charts to MS incidence and prevalence trends and comparing results across different case definitions, this study showed no statistically significant difference in case definition robustness to changes in data quality, regardless of the control limit calculations used (i.e., performance of all case definitions were affected similarly). The moderate to high proportion of OOC observations (proportions ≥ 0.30) suggests that there may be

variability in the data quality for identifying MS cases; this should be considered when using longitudinal data for MS research and surveillance.

Our previous study on juvenile diabetes that applied 0.8*SD control limits found the proportion of OOC observations for incidence and prevalence trends ranged from 0.57 to 0.76 and 0.45 to 0.83, respectively [12]. For MS, the proportions of OOC observations for both incidence and prevalence were considerably higher when using the same method. The wider control limits of 2*SD resulted in a proportion of OOC observations similar to the proportion obtained for juvenile diabetes. This suggests the wider, traditional methods of calculating control limits may be appropriate for differentiating between random and non-random error when identifying MS cases in administrative health data. MS presents differently than juvenile diabetes, and this is reflected in the healthcare encounters identified from administrative health data. Prevalence rates are also different for both diseases: in the 2019-2020 year, the Public Health Agency of Canada's Canadian Chronic Disease Surveillance System reported a prevalence of 75 per 100,000 for juvenile diabetes versus 13 per 100,000 for MS [36]. Similar to the construction of case definitions, construction of control charts to assess data quality for ascertaining different disease should carefully consider disease presentation and clinical management (i.e., the disease criteria) as well

Table 2: Comparisons of out-of-control incidence and prevalence observations for multiple sclerosis case definitions using 0.8*SD control limits

Case definition	Incidence												
	All years (1977–2017)	ICD-8 (1977–1979)		ICD-9 (1980–2004)		ICD-9/10 (2005–2017)		ICD-8 to ICD-9 (1977–1981)		ICD-9 to ICD-9/10 (2002–2006)			
	McNemar's test	McNemar's test		McNemar's test		McNemar's test		McNemar's test		McNemar's test			
	OOC frequency (Proportion)	p-value	adj. p-value ^a	p-value	adj. p-value ^a	p-value	adj. p-value ^a	p-value	adj. p-value ^a	p-value	adj. p-value ^a		
3+ (A)	32 (0.78)	REF	REF	REF	REF	REF	REF	REF	REF	REF	REF	REF	
3+ (B)	37 (0.90)	0.18	1.00	1.00	1.00	0.45	1.00	0.48	1.00	1.00	1.00	0.50	1.00
3+ (C)	29 (0.71)	0.55	1.00	1.00	1.00	1.00	1.00	0.48	1.00	1.00	1.00	1.00	1.00
7+	36 (0.88)	0.29	1.00	1.00	1.00	0.62	1.00	0.48	1.00	1.00	1.00	1.00	1.00
7+ or 3+ (coverage)	36 (0.88)	0.29	1.00	1.00	1.00	0.62	1.00	0.48	1.00	1.00	1.00	1.00	1.00
7+ or 3+ (date)	37 (0.90)	0.27	1.00	1.00	1.00	0.13	0.91	1.00	1.00	1.00	1.00	0.50	1.00
1+ hosp or 2+ phys	32 (0.78)	1.00	1.00	1.00	1.00	0.50	1.00	0.22	1.00	1.00	1.00	0.50	1.00
1+ hosp or 5+ phys	35 (0.85)	0.55	1.00	1.00	1.00	0.72	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Case definition	Prevalence												
	All years (1975–2017)	ICD-8 (1975–1979)		ICD-9 (1980–2004)		ICD-9/10 (2005–2017)		ICD-8 to ICD-9 (1975–1981)		ICD-9 to ICD-9/10 (2002–2006)			
	McNemar's test	McNemar's test		McNemar's test		McNemar's test		McNemar's test		McNemar's test			
	OOC frequency (Proportion)	p-value	adj. p-value ^a	p-value	adj. p-value ^a	p-value	adj. p-value ^a	p-value	adj. p-value ^a	p-value	adj. p-value ^a		
3+ (A)	32 (0.74)	REF	REF	REF	REF	REF	REF	REF	REF	REF	REF	REF	REF
3+ (B)	34 (0.79)	0.72	1.00	1.00	1.00	0.13	1.00	1.00	1.00	1.00	1.00	0.50	1.00
3+ (C)	31 (0.72)	1.00	1.00	1.00	1.00	0.34	1.00	0.75	1.00	1.00	1.00	1.00	1.00
7+	36 (0.84)	0.45	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
7+ or 3+ (coverage)	36 (0.84)	0.45	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
7+ or 3+ (date)	38 (0.88)	0.18	1.00	1.00	1.00	1.00	1.00	0.50	1.00	1.00	1.00	1.00	1.00
1+ hosp or 2+ phys	42 (0.98)	0.01	0.07	1.00	1.00	0.50	1.00	0.50	1.00	1.00	1.00	1.00	1.00
1+ hosp or 5+ phys	37 (0.86)	0.27	1.00	1.00	1.00	0.29	1.00	0.22	1.00	1.00	1.00	1.00	1.00

OOC: out-of-control; SD: standard deviation.

^a p-values adjusted with Holm-Bonferroni procedure.

p-values <0.05 are in boldface font.

as how the disease criteria are represented in the available data.

Previous applications of control charts in the healthcare setting have focused on monitoring healthcare delivery and population health [14, 17, 18, 37]. Trend analysis has previously been used to assess administrative health data quality for research and surveillance [11, 38–40]. Studies that used control charts to assess data quality include an applied study about juvenile diabetes case definitions [12] and a methodological study [41]. Both studies demonstrate that control charts can be used to gain insights on data quality changes over time and its impact on case definition performance. As more research is being reported about the changes in healthcare delivery due to the COVID-19 pandemic [42–44], control charts offer a potentially promising method for future work exploring the impact of COVID-19 on administrative health data quality for chronic disease research and surveillance.

The case definition of three or more hospital or physician codes, where at least one MS code must be submitted by a neurologist had the lowest proportion of OOC observations for both incidence and prevalence trends in the main and sensitivity analyses. Of note, shadow billing claims from the Manitoba MS clinic were not available between the years from 2000 to 2010, limiting the use of MS codes from neurologists during this period. Only two of eight case definitions used

observation windows; these case definitions had among the highest proportions of OOC years for prevalence. This may suggest that accounting for physician speciality in MS case definitions reduces OOC observations and including an observation window increases OOC observations; however, it is unclear, as no statistically significant differences in OOC observations between case definitions and the reference case definition were detected. It is possible that the statistical test (McNemar's test) chosen to compare OOC observations across case definitions may not be sufficiently sensitive to detect any differences.

Patterns in OOC observations across case definitions seen in Supplementary Figures 3, 4 provide some insight on the relationship between the data quality for identifying MS cases and the observed trends. The cyclical pattern seen in panels (g) and (h) in Supplementary Figure 4 (prevalence trends) suggest that case definitions that use observation windows may be influenced by external factors. Most case definitions in Supplementary Figure 3 (incidence) have an increase in incident cases (OOC outside upper control limit) around the year 2000; in 2001 the first diagnostic criteria that incorporated MRI as a means of accelerating time to diagnosis were introduced. The decrease in incident cases during 2006–2009, as seen in panels (a)–(f) could be due to missed cases from the Manitoba MS clinic when there was no shadow billing. Other potential explanations for the proportion

Table 3: Comparisons of out-of-control incidence and prevalence observations across multiple sclerosis case definitions when using 2*SD control limits

Case definition		Incidence											
		All years (1977–2017)		ICD-8 (1977–1979)		ICD-9 (1980–2004)		ICD-9/10 (2005–2017)		ICD-8 to ICD-9 (1977–1981)		ICD-9 to ICD-9/10 (2002–2006)	
		McNemar's test		McNemar's test		McNemar's test		McNemar's test		McNemar's test		McNemar's test	
		OOC frequency (Proportion)	p-value	adj. p-value ^a	p-value	adj. p-value ^a	p-value	adj. p-value ^a	p-value	adj. p-value ^a	p-value	adj. p-value ^a	p-value
3+ (A)	23 (0.56)	REF	REF	REF	REF	REF	REF	REF	REF	REF	REF	REF	REF
3+ (B)	22 (0.54)	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.5	1.0	0.5	1.0
3+ (C)	19 (0.46)	0.4	1.0	1.0	1.0	0.4	1.0	1.0	1.0	0.6	1.0	1.0	1.0
7+	24 (0.59)	1.0	1.0	1.0	1.0	1.0	1.0	0.5	1.0	1.0	1.0	1.0	1.0
7+ or 3+ (coverage)	24 (0.59)	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
7+ or 3+ (date)	27 (0.66)	0.5	1.0	0.5	1.0	0.2	1.0	0.6	1.0	0.5	1.0	1.0	1.0
1+ hosp or 2+ phys	22 (0.54)	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.5	1.0	1.0	1.0
1+ hosp or 5+ phys	29 (0.71)	0.3	1.0	1.0	1.0	0.4	1.0	0.7	1.0	1.0	1.0	1.0	1.0

Case definition		Prevalence											
		All years (1975–2017)		ICD-8 (1975–1979)		ICD-9 (1980–2004)		ICD-9/10 (2005–2017)		ICD-8 to ICD-9 (1975–1981)		ICD-9 to ICD-9/10 (2002–2006)	
		McNemar's test		McNemar's test		McNemar's test		McNemar's test		McNemar's test		McNemar's test	
		OOC frequency (Proportion)	p-value	adj. p-value ^a	p-value	adj. p-value ^a	p-value	adj. p-value ^a	p-value	adj. p-value ^a	p-value	adj. p-value ^a	p-value
3+ (A)	21 (0.49)	REF	REF	REF	REF	REF	REF	REF	REF	REF	REF	REF	REF
3+ (B)	21 (0.49)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
3+ (C)	13 (0.30)	0.06	0.37	0.50	1.00	0.34	1.00	0.48	1.00	0.50	1.00	1.00	1.00
7+	23 (0.53)	0.77	1.00	1.00	1.00	1.00	1.00	0.45	1.00	1.00	1.00	0.50	1.00
7+ or 3+ (coverage)	22 (0.51)	1.00	1.00	1.00	1.00	1.00	1.00	0.68	1.00	1.00	1.00	0.50	1.00
7+ or 3+ (date)	27 (0.63)	0.26	1.00	1.00	1.00	0.61	1.00	0.37	1.00	0.13	0.88	1.00	1.00
1+ hosp or 2+ phys	32 (0.74)	0.04	0.26	1.00	1.00	0.12	0.73	0.07	0.52	0.25	1.00	0.63	1.00
1+ hosp or 5+ phys	30 (0.70)	0.08	0.40	0.50	1.00	0.06	0.43	0.37	1.00	0.50	1.00	0.25	1.00

OOC: out-of-control; SD: standard deviation.

^a *p*-values adjusted with Holm-Bonferroni procedure.

p-values <0.05 are in boldface font.

of OOC observations include changes in ICD versions that are used to code administrative data (see panels (a)–(c) after switch to ICD-10-CA codes in hospital records) and the 7+ or 3+ (date) case definition having different criteria pre and post 1997. Moreover, disease-modifying therapy for MS became available in Manitoba in 1997/1998 and the pilot for the Manitoba MS clinic started in 1997/1998, which could have further contributed to the extreme values seen in panel (f).

Trend control charts differentiate random and non-random process variation (e.g. changes in coding practices) [13]. Trend control charts may identify where case definition validity varies, by indicating where observations are potentially influenced by changes in data quality (i.e., non-random error). The high proportion of OOC observations when using $0.8 \times \text{SD}$ control limits demonstrates greater variance in MS trends than in juvenile diabetes trends estimated from administrative health data. Wider control limits (i.e., $2 \times \text{SD}$) may help to highlight variation due to the challenges of diagnosing of MS and variation due to other factors such as policy/coding practice changes. Future work applying trend control charts to disease estimates obtained from administrative health data may benefit from the adoption of disease-specific calibrated control limits.

Acknowledgements

The authors acknowledge the Manitoba Centre for Health Policy for use of data contained in the Manitoba Population Research Data Repository under Project No. 2020-045. The results and conclusions are those of the authors and no official endorsement by the Manitoba Centre for Health Policy, Manitoba Health, or other data providers is intended or should be inferred. Data used in this study are from the Manitoba Population Research Data Repository housed at the Manitoba Centre for Health Policy, University of Manitoba and were derived from data provided by Manitoba Health.

The authors would also like to acknowledge Angela Tan, who was the analyst at MCHP for this work.

Conflicts of interest

RAM receives research funding from: Canadian Institutes of Health Research (CIHR), Research Manitoba, Multiple Sclerosis Society of Canada, Multiple Sclerosis Scientific Foundation, Crohn's and Colitis Canada, National Multiple Sclerosis Society, CMSC and the US Department of Defense, and is a co-investigator on studies receiving funding from Biogen Idec and Roche Canada. LML is supported by funding from the Canada Research Chairs program and CIHR

The other authors have no conflicts of interest to declare.

Ethics

Ethics approval was granted by the University of Manitoba Health Research Ethics Board (File No. HS23961). The Government of Manitoba's Provincial Health Research Privacy

Committee also reviewed and approved the use of Repository data for this study (PHRPC No. 2020/2021– 12).

References

- Mathai SC, Mathew S. Breathing (and Coding?) a bit easier: Changes to International Classification of Disease Coding for pulmonary hypertension. *Chest* 2018;154:207–18. <https://doi.org/10.1016/j.chest.2018.04.004>
- Nilson F, Bonander C, Andersson R. The effect of the transition from the ninth to the tenth revision of the International Classification of Diseases on external cause registration of injury morbidity in Sweden. *Inj Prev J Int Soc Child Adolesc Inj Prev* 2015;21:189–94. <https://doi.org/10.1136/injuryprev-2014-041337>
- Lavergne MR, Law MR, Peterson S, Garrison S, Hurley J, Cheng L, et al. Effect of incentive payments on chronic disease management and health services use in British Columbia, Canada: Interrupted time series analysis. *Health Policy Amst Neth* 2018;122:157–64. <https://doi.org/10.1016/j.healthpol.2017.11.001>
- Hamad AF, Vasylyk V, Yan L, Sanusi R, Ayilara O, Delaney JA, et al. Mapping three versions of the International Classification of Diseases to categories of chronic conditions. *Int J Popul Data Sci* 2021;6. <https://doi.org/10.23889/ijpds.v6i1.1406>
- Centers for Disease Control and Prevention. ICD-10 - International Classification of Diseases, Tenth Revision 2021. <https://www.cdc.gov/nchs/icd/icd10.htm> (accessed November 14, 2023).
- Lix L, Reimer K. The Canadian Chronic Disease Surveillance System: A distributed surveillance model. *Online J Public Health Inform* 2017;9. <https://doi.org/10.5210/ojphi.v9i1.7726>
- 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.
- Goodman RA, Posner SF, Huang ES, Parekh AK, Koh HK. Defining and measuring chronic conditions: Imperatives for research, policy, program, and practice. *Prev Chronic Dis* 2013;10:E66. <https://doi.org/10.5888/pcd10.120239>
- Lix L, Yogendran M, Burchill C, Charles, Metge C, McKeen N, Moore D, et al. Defining and validating chronic diseases: An administrative data approach. Winnipeg, MB: Manitoba Centre for Health Policy; 2006.
- Li L, Binney LE, Luengo-Fernandez R, Silver LE, Rothwell PM, Oxford Vascular Study. Temporal trends in the accuracy of hospital diagnostic coding for identifying acute stroke: A population-based study. *Eur Stroke J* 2020;5:26–35. <https://doi.org/10.1177/2396987319881017>

11. Hamm NC, Pelletier L, Ellison J, Tennenhouse L, Reimer K, Paterson JM, et al. Trends in chronic disease incidence rates from the Canadian Chronic Disease Surveillance System. *Health Promot Chronic Dis Prev Can Res Policy Pract* 2019;39:216–24. <https://doi.org/10.24095/hpcdp.39.6/7.02>
12. Hamm NC, Jiang D, Marrie RA, Irani P, Lix LM. Control charts for chronic disease surveillance: Testing algorithm sensitivity to changes in data coding. *BMC Public Health* 2022;22:406. <https://doi.org/10.1186/s12889-021-12328-w>
13. Sellick JA. The use of statistical process control charts in hospital epidemiology. *Infect Control Hosp Epidemiol* 1993;14:649–56. <https://doi.org/10.1086/646659>
14. Cook DA, Steiner SH, Cook RJ, Farewell VT, Morton AP. Monitoring the evolutionary process of quality: Risk-adjusted charting to track outcomes in intensive care. *Crit Care Med* 2003;31:1676–82. <https://doi.org/10.1097/01.CCM.0000065273.63224.A8>
15. Coory M, Duckett S, Sketcher-Baker K. Using control charts to monitor quality of hospital care with administrative data. *Int J Qual Health Care* 2008;20:31–9. <https://doi.org/10.1093/intqhc/mzm060>
16. Gustafson TL. Practical risk-adjusted quality control charts for infection control. *Am J Infect Control* 2000;28:406–14. <https://doi.org/10.1067/mic.2000.109883>
17. Dignam T, Hodge J, Chuks S, Mercado C, Ettinger AS, Flanders WD. Use of the CUSUM and Shewhart control chart methods to identify changes of public health significance using childhood blood lead surveillance data. *Environ Epidemiol Phila Pa* 2020;4. <https://doi.org/10.1097/EE9.0000000000000090>
18. Bundle N, Verlander NQ, Morbey R, Edeghere O, Balasegaram S, de Lusignan S, et al. Monitoring epidemiological trends in back to school asthma among preschool and school-aged children using real-time syndromic surveillance in England, 2012–2016. *J Epidemiol Community Health* 2019;73:825–31. <https://doi.org/10.1136/jech-2018-211936>
19. NHS Institute for Innovation and Improvement. A guide to creating and interpreting run and control charts: Turning data into information for improvement. Coventry: University of Warwick; 2009.
20. Marrie RA, McKay K. Administrative data for observational research in multiple sclerosis: Opportunities and challenges. *Mult Scler J* 2022;28:3–6. <https://doi.org/10.1177/13524585211055787>
21. Yamout BI, Alroughani R. Multiple Sclerosis. *Semin Neurol* 2018;38:212–25. <https://doi.org/10.1055/s-0038-1649502>
22. Katz A, Enns J, Smith M, Burchill C, Turner K, Towns D. Population data centre profile: The Manitoba Centre for Health Policy. *Int J Popul Data Sci* n.d.;4:1131. <https://doi.org/10.23889/ijpds.v5i1.1131>
23. Marrie RA, Yu N, Blanchard J, Leung S, Elliott L. The rising prevalence and changing age distribution of multiple sclerosis in Manitoba. *Neurology* 2010;74:465–71. <https://doi.org/10.1212/WNL.0b013e3181cf6ec0>
24. Iljicsov A, Milanovich D, Ajtay A, Oberfrank F, Bálint M, Dobi B, et al. Incidence and prevalence of multiple sclerosis in Hungary based on record linkage of nationwide multiple healthcare administrative data. *PloS One* 2020;15:e0236432. <https://doi.org/10.1371/journal.pone.0236432>
25. Marrie RA, Fisk JD, Stadnyk KJ, Yu BN, Tremlett H, Wolfson C, et al. The incidence and prevalence of multiple sclerosis in Nova Scotia, Canada. *Can J Neurol Sci* 2013;40:824–31.
26. Marrie RA, Patten SB, Tremlett H, Wolfson C, Warren S, Svenson LW, et al. Sex differences in comorbidity at diagnosis of multiple sclerosis: A population-based study. *Neurology* 2016;86:1279–86. <https://doi.org/10.1212/WNL.0000000000002481>
27. Schmedt N, Khil L, Berger K, Riedel O. Incidence of multiple sclerosis in Germany: A cohort study applying different case definitions based on claims data. *Neuroepidemiology* 2017;49:91–8. <https://doi.org/10.1159/000481990>
28. Al-Sakran LH, Marrie RA, Blackburn DF, Knox KB, Evans CD. Establishing the incidence and prevalence of multiple sclerosis in Saskatchewan. *Can J Neurol Sci J Can Sci Neurol* 2018;45:295–303. <https://doi.org/10.1017/cjn.2017.301>
29. Murley C, Friberg E, Hillert J, Alexanderson K, Yang F. Validation of multiple sclerosis diagnoses in the Swedish National Patient Register. *Eur J Epidemiol* 2019;34:1161–9. <https://doi.org/10.1007/s10654-019-00558-7>
30. Harrell F. Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis. 2nd ed. Springer International Publishing; 2015.
31. Cohen J. Statistical power analysis for the behavioral sciences. Academic Press; 2013.
32. Pembury Smith MQR, Ruxton GD. Effective use of the McNemar test. *Behav Ecol Sociobiol* 2020;74:133. <https://doi.org/10.1007/s00265-020-02916-y>
33. Eichstaedt KE, Kovatch K, Maroof DA. A less conservative method to adjust for familywise error rate in neuropsychological research: The Holm's sequential Bonferroni procedure. *NeuroRehabilitation* 2013;32:693–6. <https://doi.org/10.3233/NRE-130893>

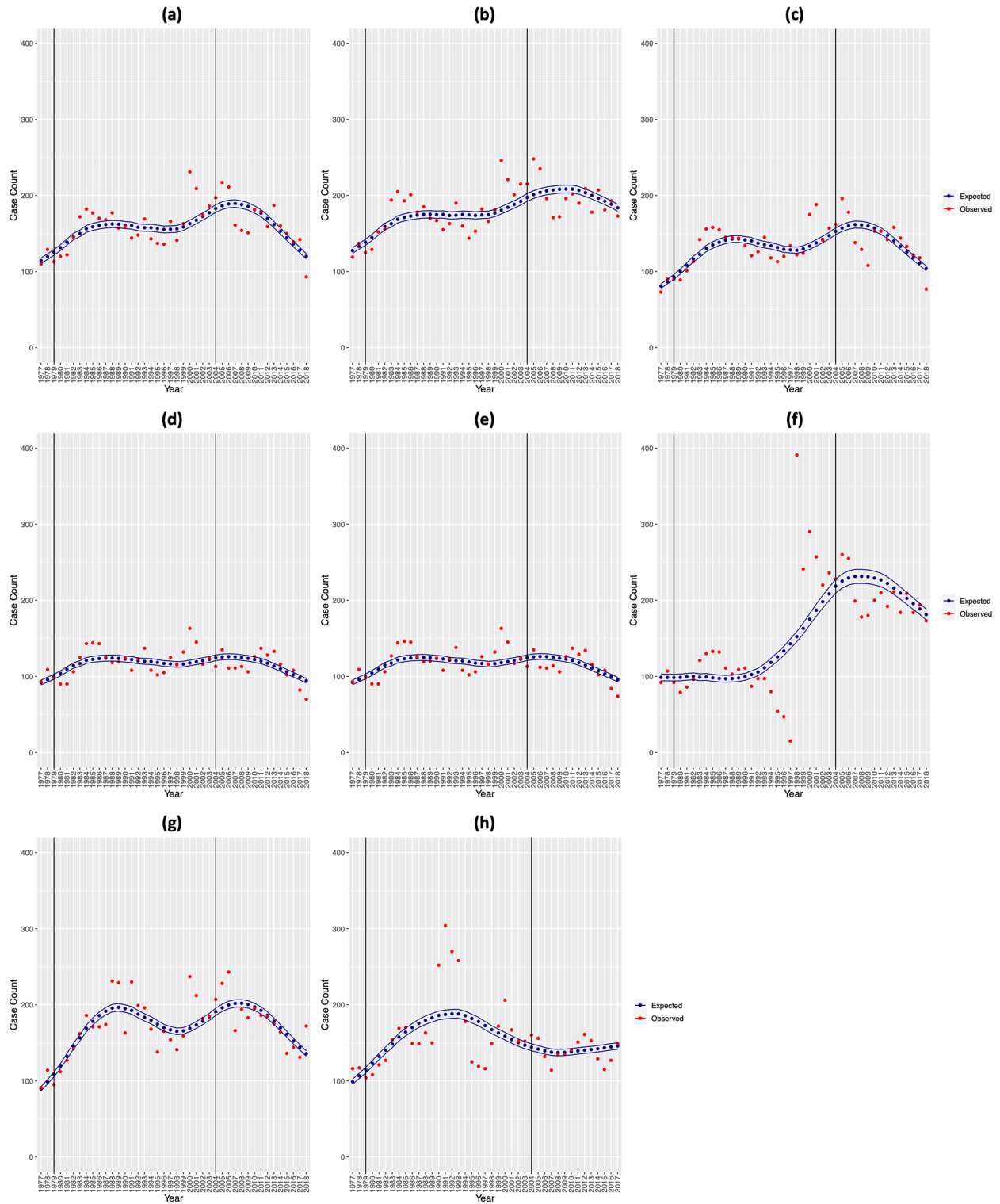
34. Venables WN, Ripley BD. Modern applied statistics with S. 4th ed. New York: Springer-Verlag; 2002. <https://doi.org/10.1007/978-0-387-21706-2>
35. Højsgaard S, Halekoh U, Yan J. The R Package geepack for Generalized Estimating Equations. J Stat Softw 2005;15:1–11. <https://doi.org/10.18637/jss.v015.i02>
36. Government of Canada. Canadian Chronic Disease Surveillance System (CCDSS) 2023. <https://health-infobase.canada.ca/ccdss/data-tool/> (accessed November 15, 2023).
37. Cockings JG, Cook DA, Iqbal RK. Process monitoring in intensive care with the use of cumulative expected minus observed mortality and risk-adjusted p charts. Crit Care 2006;10:R28. <https://doi.org/10.1186/cc3996>
38. O'Donnell S, Vanderloo S, McRae L, Onysko J, Patten SB, Pelletier L. Comparison of the estimated prevalence of mood and/or anxiety disorders in Canada between self-report and administrative data. Epidemiol Psychiatr Sci 2016;25:360–9. <https://doi.org/10.1017/S2045796015000463>
39. Kendzerska T, Sadatsafavi M, Aaron SD, To TM, Loughheed MD, FitzGerald JM, et al. Concurrent physician-diagnosed asthma and chronic obstructive pulmonary disease: A population study of prevalence, incidence and mortality. PloS One 2017;12:e0173830. <https://doi.org/10.1371/journal.pone.0173830>
40. O'Donnell S, Canadian Chronic Disease Surveillance System (CCDSS) Osteoporosis Working Group. Use of administrative data for national surveillance of osteoporosis and related fractures in Canada: Results from a feasibility study. Arch Osteoporos 2013;8:143. <https://doi.org/10.1007/s11657-013-0143-2>
41. Sanusi RA, Yan L, Hamad AF, Ayilara OF, Vasylyk V, Jozani MJ, et al. Transitions between versions of the International Classification of Diseases and chronic disease prevalence estimates from administrative health data: A population-based study. BMC Public Health 2022;22:701. <https://doi.org/10.1186/s12889-022-13118-8>
42. Papp E, Hammerich R, Saccomanno J, Sgarbossa T, Pohrt A, Schmidt B, et al. Impact of coronavirus disease 2019 on hospital admissions, health status, and behavioral changes of patients with COPD. Chronic Obstr Pulm Dis Miami Fla 2023;10:211–23. <https://doi.org/10.15326/jcopdf.2022.0383>
43. Chami N, Shah HA, Nastos S, Shaikh S, Tenenbein PK, Loughheed T, et al. Association between virtual primary care and emergency department use during the first year of the COVID-19 pandemic in Ontario, Canada. CMAJ Can Med Assoc J J Assoc Medicale Can 2023;195:E108–14. <https://doi.org/10.1503/cmaj.212051>
44. Bronskill SE, MacLagan LC, Maxwell CJ, Iaboni A, Jaakkimainen RL, Marras C, et al. Trends in health service use for canadian adults with dementia and parkinson disease during the first wave of the COVID-19 pandemic. JAMA Health Forum 2022;3:e214599. <https://doi.org/10.1001/jamahealthforum.2021.4599>

Abbreviations

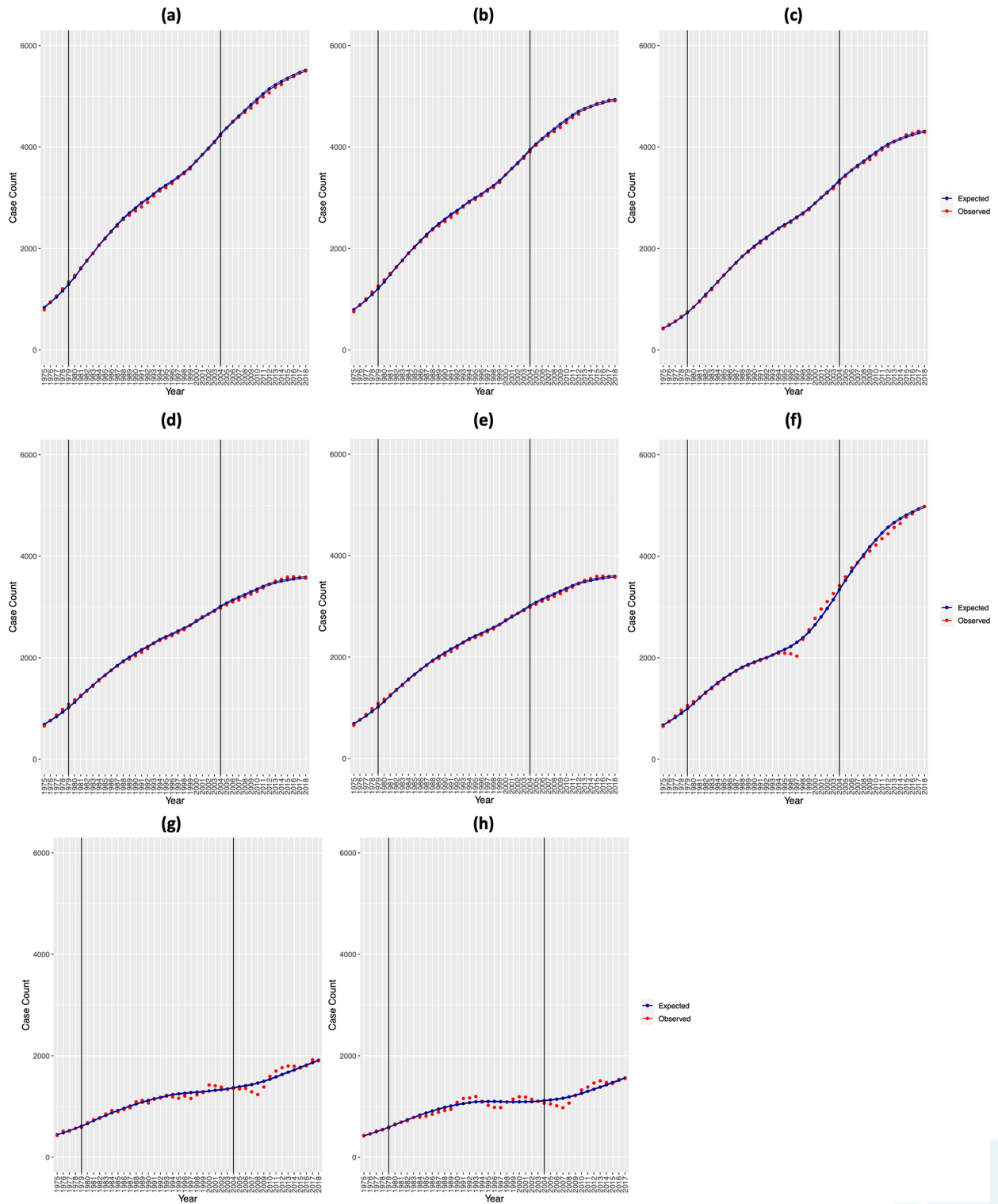
GEE:	Generalized Estimating Equations
ICD:	International Classification of Disease Codes
ICD-9-CM:	International Classification of Disease Codes-9-Clinical Modifications
ICD-10-CA:	International Classification of Disease Codes-10-Canadian Version
MRI:	Magnetic Resonance Imaging
MS:	Multiple Sclerosis
OOC:	Out-of-control



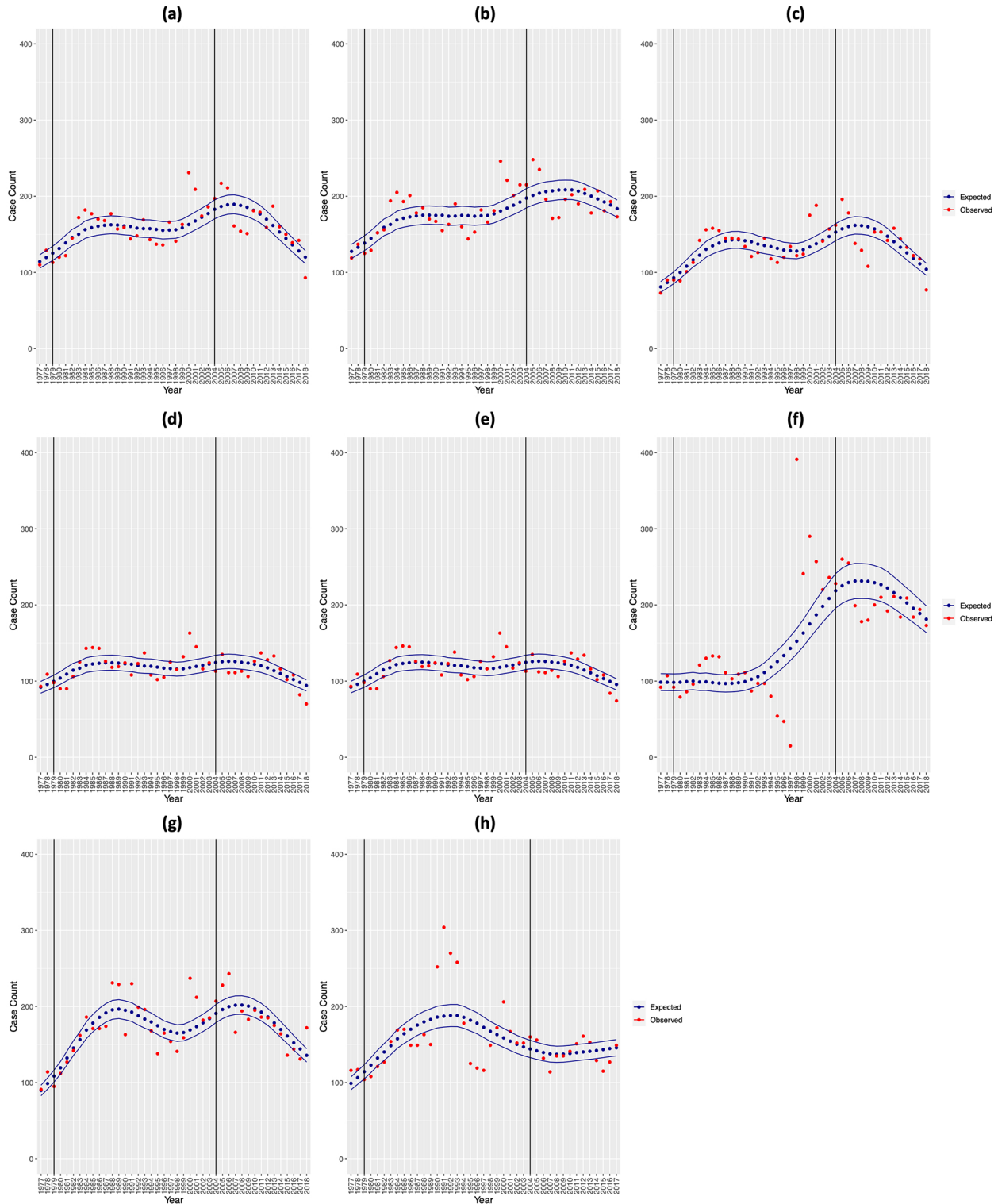
Supplementary Figure 1: Observed-expected control charts for incidence trends across multiple sclerosis case definitions using $0.8 \times \text{SD}$ control limits. (a) 3+ (A); (b) 3+ (B); (c) 3+ (C); (d) 7+; (e) 7+ or 3+ (coverage); (f) 7+ or 3+ (date); (g) 1+ hosp or 2+ phys; (h) 1+ hosp or 5+ phys. Vertical lines indicate years where a change in ICD version occurred; SD = standard deviation.



Supplementary Figure 2: Observed-expected control charts for prevalence trends for multiple sclerosis case definitions using $0.8 \times \text{SD}$ control limits. (a) 3+ (A); (b) 3+ (B); (c) 3+ (C); (d) 7+; (e) 7+ or 3+ (coverage); (f) 7+ or 3+ (date); (g) 1+ hosp or 2+ phys; (h) 1+ hosp or 5+ phys. Vertical lines indicate years where a change in ICD version occurred; SD = standard deviation.



Supplementary Figure 3: Observed-expected control charts for incidence trends across multiple sclerosis case definitions using 2*SD control limits. (a) 3+ (A); (b) 3+ (B); (c) 3+ (C); (d) 7+; (e) 7+ or 3+ (coverage); (f) 7+ or 3+ (date); (g) 1+ hosp or 2+ phys; (h) 1+ hosp or 5+ phys. Vertical lines indicate years where a change in ICD version occurred; SD = standard deviation.



Supplementary Figure 4: Observed-expected control charts for prevalence trends across multiple sclerosis case definitions using 2*SD control limits. (a) 3+ (A); (b) 3+ (B); (c) 3+ (C); (d) 7+; (e) 7+ or 3+ (coverage); (f) 7+ or 3+ (date); (g) 1+ hosp or 2+ phys; (h) 1+ hosp or 5+ phys. Vertical lines indicate years where a change in ICD version occurred; SD = standard deviation.

