BMJ Open

Building and validating trend-based multiple sclerosis case definitions: a population-based cohort study for Manitoba, Canada

Naomi C Hamm ,[1] Ruth Ann Marrie ,[1,2] Depeng Jiang ,[1] Pourang Irani,[3] Lisa Lix[1]

To cite: Hamm NC, Marrie RA, Jiang D, *et al*. Building and validating trend-based multiple sclerosis case definitions: a population-based cohort study for Manitoba, Canada. *BMJ Open* 2024;**14**:e083141. doi:10.1136/bmjopen-2023-083141

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2023-083141).

Received 13 December 2023
Accepted 01 July 2024

[1]Department of Community Health Sciences, University of Manitoba, Max Rady College of Medicine, Winnipeg, Manitoba, Canada
[2]Department of Internal Medicine, University of Manitoba, Max Rady College of Medicine, Winnipeg, Manitoba, Canada
[3]Department of Computer Science, Mathematics, Physics and Statistics, University of British Columbia, Kelowna, British Columbia, Canada

Correspondence to
Naomi C Hamm;
lettn@myumanitoba.ca

ABSTRACT

Objective This study aims to (1) build and validate model-based case definitions for multiple sclerosis (MS) that use trends (ie, trend-based case definitions) and (2) to apply dynamic classification to identify the average number of data years needed for classification (ie, average trend needed).

Design Retrospective cohort study design.

Participants 608 MS cases and 59 620 MS non-cases.

Setting Data from 1 April 2004 to 31 March 2022 were obtained from the Manitoba Population Research Data Repository. MS case status was ascertained from homecare records and linked to health data. Trend-based case definitions were constructed using multivariate generalised linear mixed models applied to annual numbers of general and specialist physician visits, hospitalisations and MS healthcare contacts or medication dispensations. Dynamic classification, which ascertains cases and non-cases annually, was used to estimate mean classification time. Classification accuracy performance measures, including sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), proportion correctly classified (PCC) and F1-scores, were compared for trend-based case definitions and a deterministic case definition of 3+MS healthcare contacts or medication dispensations.

Results When applied to the full study period, classification accuracy performance measure estimates for all case definitions exceeded 0.90, except sensitivity and PPV for the trend-based dynamic case definition (0.88, 0.64, respectively). PCC was high for all case definitions (0.94–0.99); F1-scores were lower for the trend-based case definitions compared with the deterministic case definition (0.74–0.93 vs 0.96). Dynamic classification identified 5 years as the average trend needed. When applied to the average trend windows, accuracy estimates for trend-based case definitions were lower than the estimates from the full study period (sensitivity: 0.77–0.89; specificity: 0.90–0.97; PPV: 0.54–0.81; NPV: 0.97–0.99; F1-score: 0.64–0.84). Accuracy estimates for the deterministic case definition remained high, except sensitivity (0.42–0.80). F1-score was variable (0.59–0.89).

Conclusions Trend-based and deterministic case definitions classifications were similar to a population-based clinician assessment reference standard for multiple measures of classification accuracy. However, accuracy estimates for both trend-based and deterministic case definitions varied as the years of data used for classification were reduced. Dynamic classification appears to be a viable option for identifying the average trend needed for trend-based case definitions.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This population-based cohort study assessed the performance of trend-based case definitions using data from 2004 to 2022 with a near complete capture of healthcare use trends via Manitoba's universal healthcare system.
⇒ Variations in case definition performance are identified based on the number of data years used for classification, providing applicable insights into the future application of these case definitions.
⇒ Dynamic classification provides an empirical approach to identify the average number of data years needed for classification.
⇒ The reference standard for multiple sclerosis is based on clinical assessments obtained from homecare records, which may limit generalisability.

INTRODUCTION

Administrative health data, such as physician billing claims and hospital discharge abstracts, are widely used for population-based chronic disease research and surveillance. Disease cases are identified from the data using case definitions (ie, algorithms) with either deterministic or model-based (eg, probabilistic) rules.[1–3] Accurate identification of disease cases can be challenging, as administrative health data were not originally collected for these purposes.[4] Episodic diseases, such as multiple sclerosis (MS), that have periods of remission and relapse can be particularly difficult to accurately identify in administrative health data compared to chronic diseases that are not episodic in nature.[5 6] MS is disease of the central nervous system that can lead to physical and cognitive

disability over time. Canada has among the highest prevalence of MS in the world[7–9] and is, therefore, a disease of significant interest in national surveillance initiatives.

The choice of a model-based or deterministic case definition depends on several factors, including characteristics of the disease of interest. Model-based case definitions, which rely on statistical or machine learning models to estimate case probabilities,[10–15] often have better accuracy for disease identification than deterministic case definitions,[12 13 15] which are based on a fixed number and type of observations often occurring within a defined time interval to identify cases (eg, one or more hospital separation records, or five or more physician claims within 2 years). Both approaches primarily rely on cross-sectional data and do not take the temporal characteristics of an individual's health history into account. Identifying episodic disease cases from administrative health data may benefit from using healthcare use trends (ie, longitudinal data), rather than summing an individual's healthcare history at a single time point. Studies that have used longitudinal data to predict health status have primarily relied on electronic health records or clinical data.[16–22] One study used administrative health data to detect juvenile arthritis in children[16]; however, observation began at birth, which is often not feasible for adult populations.

Case definitions that use healthcare use trends (ie, trend-based case definitions) may have variable performance as the number of data years used for classification or identification changes. Moreover, the number of data years required for accurate classification may vary across individuals. Therefore, identifying the average number of years needed for accurate case ascertainment (ie, average trend needed) is a critical first step when using healthcare use trends to identify episodic disease cases within administrative health data. Dynamic classification aims to minimise observation time[19 20] and is a potential approach for identifying the average trend needed. Individuals are classified using probability intervals and estimated interval limits are updated at regular time points.[20] If the predetermined classification cut-off value falls outside an individual's estimated interval, the individual is classified; otherwise, observation continues. Therefore, classification occurs throughout the observation period and only when enough data are present to make a 'confident classification' (ie, full probability interval either above or below classification cut-off value), allowing observation time to vary across individuals.

It is unclear how the accuracy of trend-based case definitions for episodic diseases compares to the accuracy of deterministic case definitions currently used in research and surveillance. In addition, the application of dynamic classification to identify the average trend needed for classification has not yet been tested. Our study purpose was to build and validate model-based case definitions for an episodic disease, MS, that use trends (ie, trend-based case definitions) and apply dynamic classification to identify the average number of data years needed for classification (ie, average trend needed). The objectives were to (1) assess classification performance of trend-based case definitions and a previously validated deterministic case definition using MS status obtained from homecare records as a reference standard; (2) identify the average trend needed for MS case ascertainment using dynamic classification and (3) compare the classification performance of trend-based and deterministic case definition over time using the average trend needed.

METHODS

Study design

A retrospective cohort study design was used; the study period was from 1 April 2004 to 31 March 2022. Case definitions were initially applied to data from the full study period. The study period was then split into 5-year windows (fiscal year used) based on the average trend needed as identified by the dynamic classifier (average trend windows; 1 April 2004–31 March 2009; 1 April 2009–31 March 2014; 1 April 2014–31 March 2019; 1 April 2019–31 March 2022).

Patient and public involvement

None.

Data source

Data were obtained from the Manitoba Population Research Data Repository housed at the Manitoba Centre for Health Policy, University of Manitoba. Manitoba has a universal healthcare system and captures all publicly insured healthcare contacts for its 1.3 million residents. The Manitoba Health Insurance Registry, Hospital Discharge Abstracts, Medical Claims/Medical Services, Drug Program Information Network, Canada Census and Home Care Assessment databases were used. Data on health insurance coverage dates, birth date, sex and postal code were obtained from the Manitoba Health Insurance Registry; the registry was also used to conduct individual-level linkage of databases. The Hospital Discharge Abstracts (5-digit International Classification of Diseases Codes (ICD)-10-Canadian version (CA) codes), Medical Claims/Medical Services (3-digit ICD-9-Clinical Modification (CM) codes) and Drug Program Information Network (Anatomical Therapeutic Chemical (ATC) codes) databases were used to obtain information on healthcare visits and prescription medications from community pharmacies. The Canada Census was used to obtain area-level income quintile based on postal code and average household income.[23] The Home Care Assessment database was used to construct the study cohort and provide a reference standard for identifying MS cases and non-cases. This database captures data on home care assessments, utilisation and health status for all individuals receiving homecare delivered by the Winnipeg Regional Health Authority. The Winnipeg Regional Health Authority is the largest health authority

in the province and serves approximately 60% of Manitoba's population.

Reference standard

The reference standard for MS status was based on the interRAI assessment obtained from the Homecare Assessment database. The interRAI assessment is an internationally recognised tool that assesses an individual's health and functioning and has been shown to have high sensitivity (0.90) and specificity (1.00) for identifying individuals with MS within the homecare setting.[24] Assessments are completed by a clinician based on patient interviews and review of medical information[24]; within Manitoba, an assessment is required for homecare access. Indication of MS was determined from a checklist of conditions; those with MS checked were considered cases and those without were considered non-cases. If multiple assessments per individual were available, the status of the majority of assessments was used. Therefore, individuals were assigned the case status that had the strongest evidence of true MS status. Where the numbers of MS case/non-case assessments were equal, individuals were excluded. Overall, the proportion of individuals that had conflicting MS statuses in the data was low (0.002% of cohort).

Study cohort

Individuals were included in the study if they had one or more assessments in the Homecare Assessment database during the study period. Cohort entry was defined as the start of the study period (1 April 2004) or start of healthcare coverage, whichever was later. Cohort inclusion criteria were a valid MS assessment field (ie, a response from an assessment that had been signed by a physician), linkage to the Manitoba Health Insurance Registry, at least 730 days of continuous healthcare coverage between cohort entry and assessment date and at least 20 years of age at assessment date. 20 years was chosen as the minimum cut-off age as this is the age used by the Public Health Agency of Canada for MS surveillance.[25]

Study variables

Four healthcare use variables were used to build longitudinal case definitions: the number of general physician (ie, family physician) visits, the number of specialist physician visits, at least one inpatient hospitalisation for any reason and at least one MS healthcare contact (ie, physician visit or hospitalisation with an MS diagnosis code or a MS-specific prescription medication claim). These measures were constructed for each year in the study period. The number of general physician visits and the number of specialist physician visits were capped at 92 visits per year (ie, 1 visit every 4 days; <1% of cohort affected). Specialist physician visits encompassed any specialty because the MS population has a higher co-occurrence of multiple health conditions than the general population.[26 27] Neurologist visits in the Manitoba MS Clinic were not captured in administrative data between 2000 and 2010 due to a lack of shadow billing for alternate-funded physicians and were, therefore, excluded from specialist visits. MS diagnosis codes were ICD-9-CM 340 and ICD-10-CA G35.[28] ATC codes for MS-specific prescription medications are reported in online supplemental table S1.[29] Demographic variables included age at cohort entry, sex, income quintile and years of healthcare coverage during the study period.

Case definitions

Three types of case definitions were applied to the data: Trend-based with dynamic classification, where classification was based on credible intervals (CrI) calculated annually (trend-based dynamic; more details on the dynamic classification scheme are found in the Statistical Analysis section); trend-based with static classification, where classification was based on a single probability point estimate calculated from data over the full study period (trend-based static) and a previously validated deterministic case definition (3 or more MS contacts over the full study period).[28 30] The trend-based case definitions were built using group-specific multivariate generalised linear mixed models (ie, separate models for cases and non-cases). A full description of the models and methods used to calculate case probability and corresponding CrIs can be found in online supplemental material and in Hughes *et al*.[20] Outcome variables included the four healthcare use variables (general physician visits, specialist physician visits, hospitalisation and MS contacts). Count outcome variables (general and specialist physician visits) were modelled with a log link function and binary outcome variables (hospitalisation and MS contacts) were modelled with a logit link function. Model covariates included time (time=0 for the first year in the study period), sex, age at cohort entry (continuous) and income quintile (quintiles 1–3=low income; 4–5=high income); all covariates were binary or continuous and no transformations were used. All models included a random intercept. Based on univariate analyses (online supplemental table S2), general visits and specialist visits were modelled with a random time slope; hospitalisations were modelled assuming a fixed time slope. MS-specific contacts were modelled with a random intercept and no covariates due to their sparse numbers in the non-cases. The same outcome variables and covariates were used in all models (ie, case and non-case models).

Individual group probabilities were estimated by using Bayesian methods approximated via a Markov Chain Monte Carlo with a burn-in of 500 iterations and a thinning rate of 100.[20] Trace plots were used to determine burn-in rates and autocorrelation plots were used to assess thinning rates.[31] Model convergence was assessed using trace plots, Gelman-Rubin-Brooks plots and the Gelman-Rubin diagnostic; when the trace plots for all coefficients had strong overlap and the Gelman-Rubin diagnostic was <1.1, the model was considered sufficiently converged.[31 32] Trace, density, Gelman-Rubin-Brooks and

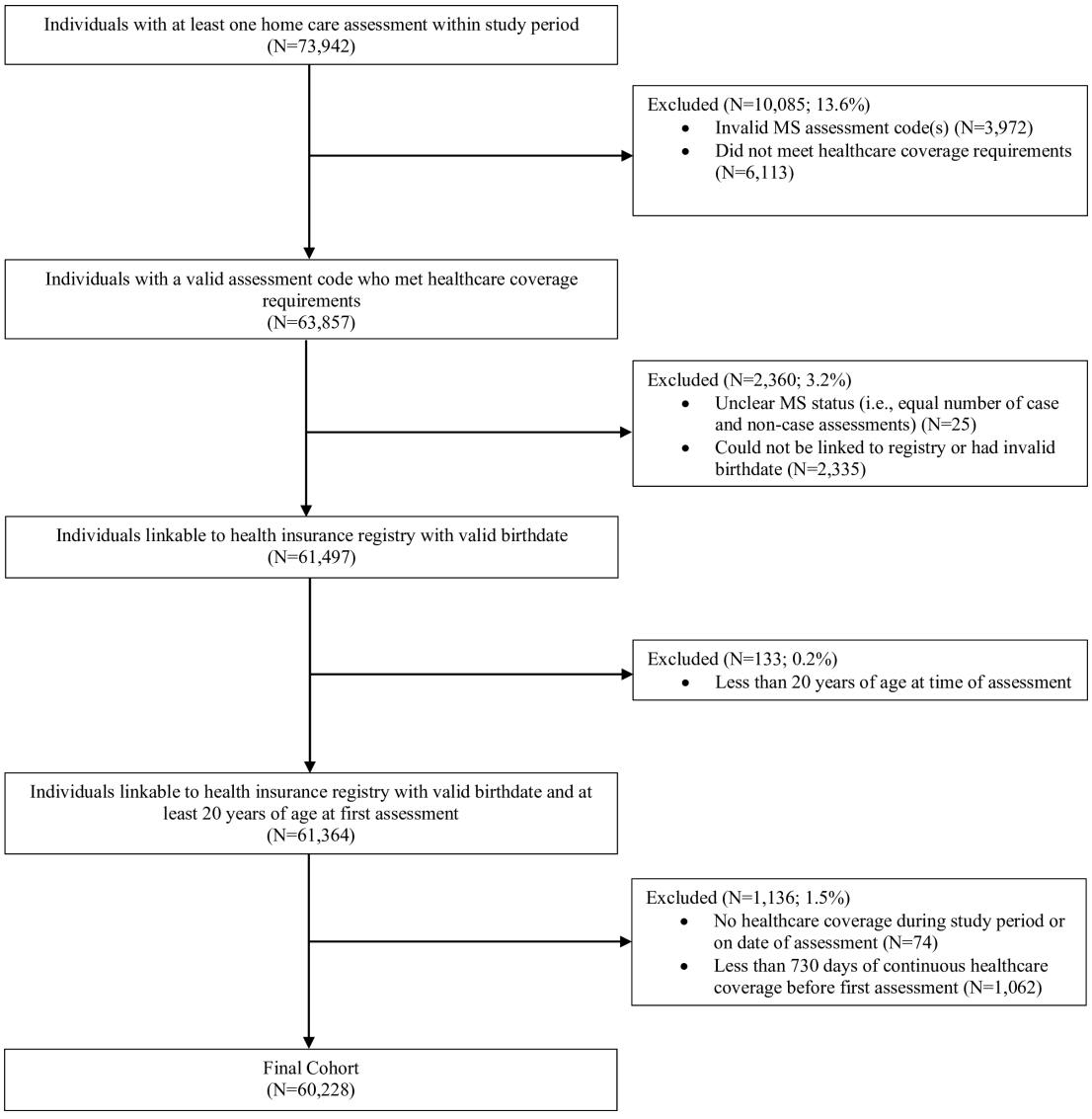

Figure 1 Flow chart for study cohort. MS, multiple sclerosis.

autocorrelation plots can be found in online supplemental figures S1 and S2.

Statistical analysis

Study cohort characteristics were described using means, SD, medians, IQRs, frequencies and percentages based on variable type. Group (ie, case vs non-case) differences were tested using Student's t-tests for continuous measures and χ^2 tests of independence for categorical measures.

Models for the trend-based case definitions were applied to data from the full study period. A classification cut-off point, denoted as c, was determined as the value nearest to the top left corner of the receiver operating characteristic (ROC) curve. For the trend-based static case definition, individual probabilities were estimated using data from the full study period. For the trend-based dynamic case definition, classification was as follows:

1. Calculate individual probabilities and their corresponding 95% CrI for year 1 of study period ($P^{LOW}(t)$, $P^{UPP}(t)$).
2. If $P^{LOW}(t) > c$, classify individual as a case.
3. If $P^{UPP}(t) < c$, classify individual as a non-case.
4. If $P^{LOW}(t) \leq c \leq P^{UPP}(t)$, leave individual unclassified.
5. If individual remains unclassified, follow to next year and update probability of corresponding CrI and repeat steps 2–5.

Using this classification scheme, individuals had different classification times (ie, the number of data years) depending on the observation year where their CrI was either fully below the cut-off point (non-case) or fully above the cut-off point (case).

Case definition performance was evaluated using the following accuracy measure estimates: area under the curve (AUC), sensitivity, specificity, positive and negative predictive values (PPV, NPV), proportion of individuals correctly classified (PCC) and F1-scores. Trend-based

Table 1 Description of cohort characteristics

Variable	Cases (n=608) Frequency (%)	Non cases (n=59 620) Frequency (%)	P value
Sex			0.0104
Males	195 (32.1)	22 130 (37.1)	
Females	413 (67.9)	37 490 (62.9)	
Period of MS assessment			<0.0001
2004–2009	331 (54.4)	22 691 (38.1)	
2010–2015	148 (24.3)	19 182 (32.2)	
2016–2021	129 (21.2)	17 747 (29.8)	
Income quintile at index date			<0.0001
Q1 (lowest)	101 (16.6)	14 245 (23.9)	
Q2	122 (20.1)	13 096 (22.0)	
Q3	125 (20.6)	12 112 (20.3)	
Q4	112 (18.4)	10 156 (17.0)	
Q5 (highest)	142 (23.4)	9542 (16.0)	
	Mean (SD)	**Mean (SD)**	
Age at cohort entry	54.2 (12.8)	69.4 (13.2)	<0.0001
Age at assessment date	61.2 (12.7)	77.9 (12.3)	<0.0001
Total healthcare coverage, years	15.0 (4.2)	13.3 (4.7)	<0.0001
Healthcare coverage before assessment date, years	7.0 (4.9)	8.5 (4.9)	<0.0001
Healthcare coverage after assessment date, years	8.0 (5.0)	4.8 (3.8)	<0.0001

MS, multiple sclerosis.

case definitions were evaluated using five-fold cross-validation. Due to the computational intensity of the models, random samples of non-cases were selected for both training and validation models (1:5 case to non-case ratio for training models; 1:10 case to non-case ratio for validation).

The average trend needed was identified as the mean classification time, in years, when applying the trend-based dynamic case definition. All case definitions were then reapplied to average trend windows and performance was reassessed.

Supplementary analyses were conducted to explore factors that may influence classification time for the trend-based dynamic case definition. After the trend-based dynamic case definition was applied, classification time for the entire cohort was split into quintiles. Quintile means and proportions of cohort characteristics were calculated and stratified by MS case status. Case definitions were also applied to an additional average trend window (1 April 2017–31 March 2022) as a supplementary analysis, as the full study period could not be evenly split into 5-year windows.

All data analyses were performed by using R V.4.1.0[33] and SAS V.9.4 (SAS Institute). The mixAK package[34] was used to build and validate the trend-based case definitions. SAS was used to apply and validate the deterministic case definition.

RESULTS

Between 1 April 2004 and 31 March 2022, 60 228 eligible individuals (608 MS cases (1.0%), 59 620 non-cases (99.0%)) were identified in the Homecare database (figure 1). Cohort characteristics are provided in table 1. Cases comprised a slightly higher percentage of females compared with non-cases (cases 68% female; non-cases 63% female). A larger proportion of cases were in the higher-income quintiles, whereas a larger proportion of non-cases were in the lower-income quintiles. At cohort entry and MS assessment date, cases had a mean age of 54 and 61 years and non-cases had a mean age of 69 and 78 years, respectively. Cases had a slightly greater average total number of years of healthcare coverage than non-cases. Non-cases had more healthcare coverage before the assessment date and less coverage after the assessment date than cases.

Table 2 reports accuracy estimates for the case definitions applied to the full study period. The trend-based static case definition had the highest sensitivity estimate (0.96, SD: 0.02) and the deterministic case definition had the highest PPV estimate (0.98, SD: 0.005); specificity and NPV estimates were similar for all three case definitions. The AUC estimate was slightly higher for the trend-based static case definition compared with the trend-based dynamic (0.98 vs 0.94). The trend-based dynamic case definition had the lowest PPV, sensitivity and F1-score

Table 2 Classification accuracy measure estimates (SEs) for trend-based and deterministic case definitions for the full study period

Accuracy measure	Trend-based dynamic case definition*	Trend-based static case definition*	Deterministic case definition
AUC	0.94	0.98	N/A
Sensitivity	0.88 (0.03)	0.96 (0.02)	0.94 (0.009)
Specificity	0.95 (0.006)	0.99 (0.003)	1.00 (0.001)
PPV	0.64 (0.04)	0.91 (0.02)	0.98 (0.005)
NPV	0.99 (0.003)	1.00 (0.002)	0.99 (0.001)
PCC	0.94	0.99	0.99
F1-Score	0.74	0.93	0.96

*Estimates and SEs are averaged over fivefold.
AUC, area under the curve; N/A, not applicable; NPV, negative predictive value; PCC, proportion correctly classified; PPV, positive predictive value.

estimates (0.64,0.88 and 0.74, respectively). The PCC estimates were similar for all case definitions. The trend-based static and deterministic case definition had similar F1-scores. Mean classification time for the trend-based dynamic case definition was 5 years; this estimate was used to define the average trend windows.

Accuracy estimates for the trend-based case definitions applied to the average trend windows were slightly lower than the accuracy estimates obtained for trend-based case definitions applied to the full study period (table 3). PPV and F1-score estimates had the lowest values, which ranged from 0.54 to 0.81 and 0.64 to 0.84, respectively. In contrast, the deterministic case definition had similar accuracy measure estimates when applied to the average trend windows and the full study period, except sensitivity and F1-score, which had considerably lower estimates for the average trend windows (sensitivity: 0.41–0.80 vs 0.94; F1-score: 0.59–0.89 vs 0.96). Variability in sensitivity, specificity, PPV and NPV estimates for the three case definitions across the average trend windows can be seen in online supplemental figure S3. Supplementary analyses for the average trend window of 1 April 2017–31 March 2022 can be found in online supplemental table S3. Classification accuracy measures for all cases were slightly higher when applied to the 2017–2022 average trend window compared with classification accuracy measures obtained using data from the 2019–2022 average trend window.

Mean case probability estimates and 95% CrIs for cases and non-cases over the study period are reported in figure 2. Mean case probability estimates increased over time for cases and decreased for non-cases. Mean 95% CrIs decreased over time for non-cases.

Results from the supplementary analyses exploring the impact of cohort characteristics on classification time are found in online supplemental tables S4-S6. For cases and non-cases, individuals who were younger at cohort entry were more likely to be classified within the first classification time quintile. For cases, a higher proportion of individuals classified within classification time quintile 1 were in income quintile 5 (Q5; 0.25) compared with the remaining income quintiles. For non-cases, the highest proportion of individuals classified within classification time quintile 1 was in quintiles 1 and 2 (Q1 and Q2; 0.24 and 0.25, respectively). The highest proportion of cases were classified in year 1 (online supplemental figure S4), whereas the highest proportion of non-cases were classified in year 5 (online supplemental figure S5).

DISCUSSION

This study aimed to build and validate trend-based case definitions for MS and assess the use of dynamic classification for identifying the average trend needed for classification. Trend-based case definition performance was compared with the performance of a previously validated deterministic case definition. We found similar accuracy estimates for trend-based dynamic, trend-based static and deterministic case definitions; the trend-based dynamic case definition had lower PPV and sensitivity estimates compared with the other case definitions. Dynamic classification estimated an average trend of 5 years was needed for classification. When the observation period was limited to the average trend needed (ie, 5 years), performance of all case definitions was slightly lower; sensitivity for the deterministic case definition was considerably lower. Poorest performance estimates were observed for all case definitions when they were applied to the most recent trend window (ie, 1 April 2019–31 March 2022).

Previous studies validating MS case definitions for administrative health data have reported variable accuracy estimates.[28 30 35–42] The majority of validated case definitions were deterministic. The deterministic case definition used in this study (three or more MS contacts) has been validated in multiple geographical regions and populations (children and adults), with good performance (sensitivity: 0.87–0.99; specificity: 0.56–1.00; PPV: 0.75–1.00; NPV: 0.76–0.98).[28 30 36–39]

The estimated PPV values obtained when applying the trend-based case definitions to the average trend windows

Table 3 Classification accuracy measure estimates and ranges for trend-based and deterministic case definitions for the average trend windows

Accuracy measure	2004–2009	2009–2014	2014–2019	2019–2022
Trend-based dynamic case definition*				
AUC	0.92	**0.94**	0.92	0.88
Sensitivity	0.85	**0.87**	0.86	0.77
Specificity	0.93	**0.95**	0.93	0.91
PPV	0.55	**0.64**	0.60	0.55
NPV	0.98	**0.99**	0.98	0.97
PCC	0.92	**0.94**	0.92	0.90
F1-Score	0.67	**0.74**	0.71	0.64
Trend-based static case definition*				
AUC	**0.94**	**0.94**	0.93	0.89
Sensitivity	0.87	**0.89**	0.88	0.79
Specificity	0.94	**0.97**	**0.97**	0.90
PPV	0.61	0.77	**0.81**	0.54
NPV	**0.99**	**0.99**	**0.99**	0.97
PCC	0.94	**0.96**	**0.96**	0.89
F1-Score	0.72	0.83	**0.84**	0.64
Deterministic case definition				
Sensitivity	0.76	**0.80**	0.68	0.42
Specificity	**1.00**	**1.00**	**1.00**	1.00
PPV	**1.00**	**1.00**	0.99	0.98
NPV	**0.98**	**0.98**	0.97	0.94
PCC	**0.98**	**0.98**	0.97	0.95
F1-Score	0.86	**0.89**	0.81	0.59

Boldface indicates maximum estimate observed over the average trend windows.
Underline indicates minimum estimate observed over the average trend windows.
*Estimates are averaged over fivefold.
AUC, area under the curve; NPV, negative predictive value; PCC, proportion correctly classified; PPV, positive predictive value.

were lower than the PPV estimates obtained when applying trend-based case definitions to the full study period. In contrast, the deterministic case definition had lower sensitivity estimates when applied to the average trend windows compared with when it was applied to the full study period. This indicates trend-based case definitions are more likely to misclassify non-cases, whereas the deterministic case definition is more likely to misclassify cases when the number of data years used for classification is reduced. Understanding how case definitions are robust to changes in the years of data used for classification is important when applying case definitions across multiple jurisdictions, as most jurisdiction do not have the same number of data years available.[43 44]

As trend-based case definitions rely on trends for classification, changes in disease treatment over time may influence performance. While this applies to the full study period, it is most evident when applying case definitions to the 2019–2022 average trend window, where drastic changes in healthcare were observed due to the COVID-19 pandemic, including increased virtual care, physician departure from clinics and introduction of new MS medications.[45 46] Data exploration indicated a drop in the mean number of physician visits (general and specialist), hospitalisations and MS-specific contacts for the 2021 fiscal year due to the COVID-19 pandemic, which likely contributed to lower estimates of case definition performance. Lower estimates were still observed when the average trend window was extended to 1 April 2017–31 Marc 2022 in supplementary analysis.

As expected, mean case probability estimates for cases increased over the study period, whereas mean case probability estimates decreased for non-cases. This indicates that using a longer trend for classification (ie, more years of data) resulted in a more accurate estimated case probability. Notably, the estimated case probability trend seen in this study was based on a classification approach and obtained under the assumption that MS prevalence remained constant over time, which is not always the case.[28 47] Worldwide, MS prevalence is considered to be increasing primarily due to earlier diagnosis and improved survival rates.[47] A different trend in estimated

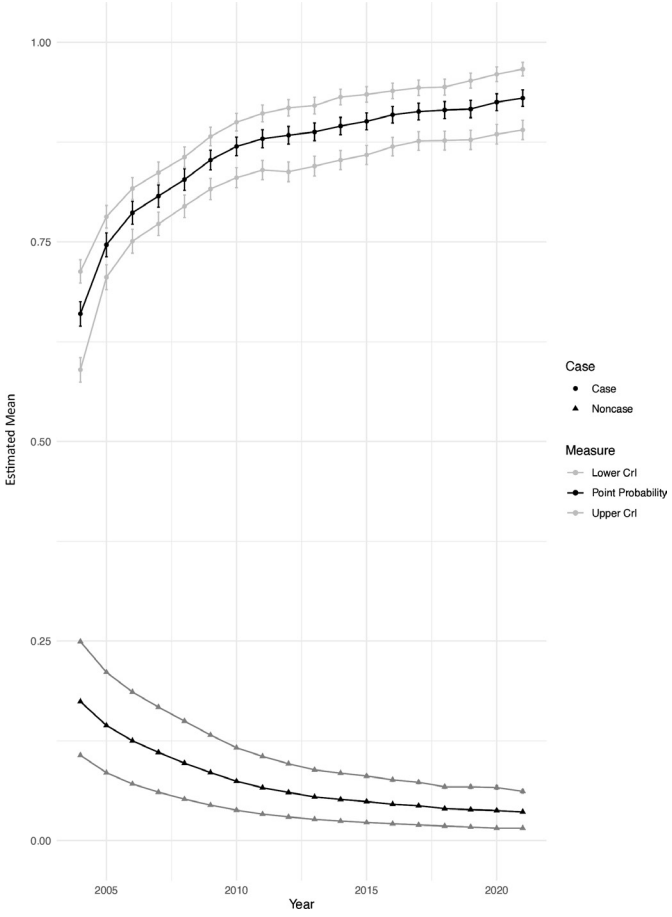

Figure 2 Mean estimated case probability and 95% credible interval (CrI) limits with SE bars at each year during observation period for cases and non-cases.

case probability may be observed when changes in baseline prevalence are considered or where a prediction, rather than classification approach is used.

There were some limitations to this study. The selected reference standard for MS status and study cohort comes from those receiving homecare within the largest health authority in the province of Manitoba, which primarily serves urban populations. Therefore, study finding may not generalise to younger, healthier populations or rural populations. MS status was obtained from the InterRAI assessment. This is a validated assessment conducted by a clinician[24 48]; however, MS cases and non-cases may still be misclassified.[24] The interRAI assessment has been previously used as a validation source for MS.[36]

Strengths of the study include near-complete capture (ie, 99%) of healthcare use via Manitoba's universal healthcare system.[49–51] In addition, applying case definitions to a reduced number of data years as well as the full study period provides a more complete picture of case definition performance. The deterministic case definition chosen for this study as a comparison for the trend-based case definitions has been well validated in Manitoba[27 28 52] as well as other jurisdictions.[30 38 52] Last, the novel use of dynamic classification provides an empirical and effective approach to identify the average trend needed, which can easily be applied to different episodic disease in future research.

In conclusion, trend-based case definitions have similar performance to deterministic case definitions when identifying MS cases and non-cases from administrative health data. Performance for both trend-based and deterministic case definition varies when the number of data years used for classification is limited. When using a trend-based case definition, dynamic classification appears to be a viable option for identifying the average trend needed for classification. Future research should examine how changes in the years of data used for classification at the individual-level impact case definition performance, as we only explored changes in the number of data years at the population (ie, marginal) level. The application of trend-based case definitions should also be explored for other episodic chronic diseases, such as rheumatoid arthritis[53 54] or inflammatory bowel disease.[55 56]

Contributors NCH and LL conceived the idea for the study. NCH, LL, DJ, RAM and PI defined the scope of the study and created the analysis plans. NCH conducted the analyses and had access to individual-level data obtained within study period. NCH and LL drafted the manuscript and DJ, RAM and PI contributed to its revisions. NCH, LL, DJ, RAM and PI reviewed and approved the final manuscript for submission. NCH is the guarantor.

Funding This work was supported by the Canadian Institutes of Health Research [FDN-143293]. NCH received funding from the Visual and Automated Disease Analytics Trainee ProgramProgramme during the time of this study. LL is supported by a Canada Research Chair in Methods for Electronic Health Data Quality (CRC-2017–00186). PI is supported by a Canada Research Chair in Ubiquitous Analytics. RAM is supported by the Waugh Family Chair in Multiple Sclerosis and a Manitoba Research Chair from Research Manitoba.

Competing interests RAM receives research funding from: 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. LL receives research funding from CIHR, Canada Research Chairs Programme, Natural Sciences and Engineering Research Council of Canada, National Institutes of Health, and the Canadian Agency for Drugs and Technologies in Health.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Ethics approval was granted by the University of Manitoba's Health Research Ethics Board (HREB No. HS23961). Data access approval was provided by the Provincial Health Research Privacy Committee (PHRPC No. 2020/2021-12) and Manitoba Shared Health along with the Winnipeg Regional Health Authority (RAAC2020:026).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Data used in this article were derived from administrative health data as secondary use. The data were provided to the investigators under specific data sharing agreements only for approved use at Manitoba Centre for Health Policy (MCHP). The original source data are not owned by the researchers or MCHP and as such cannot be provided to a public repository. The original data source and approval for use have been noted in the Ethics section of the article. Where necessary, source data specific to this article or project may be reviewed at MCHP with the consent of the original data providers, along with the required privacy and ethical review bodies.

ORCID iDs
Naomi C Hamm http://orcid.org/0000-0002-8646-8995
Ruth Ann Marrie http://orcid.org/0000-0002-1855-5595
Depeng Jiang http://orcid.org/0000-0002-7060-4892

REFERENCES

1 Lix LM, Ayles J, Bartholomew S, *et al*. The Canadian chronic disease surveillance system: a model for collaborative surveillance. *Int J Popul Data Sci* 2018;3:433.
2 Lix LM, Yogendran MS, Shaw SY, *et al*. Population-based data sources for chronic disease surveillance. *Chronic Dis Can* 2008;29:31–8.
3 Rector TS, Wickstrom SL, Shah M, *et al*. Specificity and sensitivity of claims-based algorithms for identifying members of medicare choice health plans that have chronic medical conditions. *Health Serv Res* 2004;39:1839–57.
4 Iezzoni LI. Assessing quality using administrative data. *Ann Intern Med* 1997;127:666–74.
5 Marrie RA, McKay K. Administrative data for observational research in multiple sclerosis: opportunities and challenges. *Mult Scler* 2022;28:3–6.
6 Yamout BI, Alroughani R. Multiple sclerosis. *Semin Neurol* 2018;38:212–25.
7 Petrin J, Marrie RA, Devonshire V, *et al*. Good multiple sclerosis (MS) care and how to get there in Canada: perspectives of Canadian healthcare providers working with persons with MS. *Front Neurol* 2023;14:1101521.
8 GBD 2016 Multiple Sclerosis Collaborators. Global, regional, and national burden of multiple sclerosis 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* 2019;18:269–85.
9 Amankwah N, Marrie RA, Bancej C, *et al*. Multiple sclerosis in Canada 2011 to 2031: results of a microsimulation modelling study of epidemiological and economic impacts. *Health Promot Chronic Dis Prev Can* 2017;37:37–48.
10 Cooke CR, Joo MJ, Anderson SM, *et al*. The validity of using ICD-9 codes and pharmacy records to identify patients with chronic obstructive pulmonary disease. *BMC Health Serv Res* 2011;11:37.
11 English SW, McIntyre L, Fergusson D, *et al*. Subarachnoid hemorrhage admissions retrospectively identified using a prediction model. *Neurology (ECronicon)* 2016;87:1557–64.
12 Fan J, Arruda-Olson AM, Leibson CL, *et al*. Billing code algorithms to identify cases of peripheral artery disease from administrative data. *J Am Med Inform Assoc* 2013;20:e349–54.
13 Lix LM, Yogendran MS, Leslie WD, *et al*. Using multiple data features improved the validity of osteoporosis case ascertainment from administrative databases. *J Clin Epidemiol* 2008;61:1250–60.
14 Peng M, Chen G, Lix LM, *et al*. Refining hypertension surveillance to account for potentially misclassified cases. *PLoS One* 2015;10:e0119186.
15 Prosser RJ, Carleton BC, Smith MA. Identifying persons with treated asthma using administrative data via latent class modelling. *Health Serv Res* 2008;43:733–54.
16 Feely A, Lim LS, Jiang D, *et al*. A population-based study to develop juvenile arthritis case definitions for administrative health data using model-based dynamic classification. *BMC Med Res Methodol* 2021;21:105.
17 Choi E, Schuetz A, Stewart WF, *et al*. Using recurrent neural network models for early detection of heart failure onset. *J Am Med Inform Assoc* 2017;24:361–70.
18 Reddy BK, Delen D. Predicting hospital readmission for lupus patients: an RNN-LSTM-based deep-learning methodology. *Comput Biol Med* 2018;101:199–209.
19 Hughes DM, Komárek A, Czanner G, *et al*. Dynamic longitudinal discriminant analysis using multiple longitudinal markers of different types. *Stat Methods Med Res* 2018;27:2060–80.
20 Hughes DM, Komárek A, Bonnett LJ, *et al*. Dynamic classification using credible intervals in longitudinal discriminant analysis. *Stat Med* 2017;36:3858–74.
21 Wang T, Qiu RG, Yu M. Predictive modeling of the progression of Alzheimer's disease with recurrent neural networks. *Sci Rep* 2018;8:9161.
22 Maruyama N, Takahashi F, Takeuchi M. Prediction of an outcome using trajectories estimated from a linear mixed model. *J Biopharm Stat* 2009;19:779–90.
23 Martens P, Nickel N, Forget E, *et al*. *The cost of smoking: a Manitoba study*. Winnipeg, MB: Manitob Centre for Health Policy, 2015. Available: http://mchp-appserv.cpe.umanitoba.ca/deliverablesList.html
24 Foebel AD, Hirdes JP, Heckman GA, *et al*. Diagnostic data for neurological conditions in interRAI assessments in home care, nursing home and mental health care settings: a validity study. *BMC Health Serv Res* 2013;13:457.
25 Government of Canada. Canadian chronic disease surveillance system (CCDSS) data tool. 2023. Available: https://health-infobase.canada.ca/ccdss/data-tool/Index
26 Marrie RA, Patten SB, Tremlett H, *et al*. Sex differences in comorbidity at diagnosis of multiple sclerosis: a population-based study. *Neurology (ECronicon)* 2016;86:1279–86.
27 Marrie RA, Fisk JD, Yu BN, *et al*. Mental comorbidity and multiple sclerosis: validating administrative data to support population-based surveillance. *BMC Neurol* 2013;13:16.
28 Marrie RA, Yu N, Blanchard J, *et al*. The rising prevalence and changing age distribution of multiple sclerosis in Manitoba. *Neurology (ECronicon)* 2010;74:465–71.
29 Marrie RA, Kosowan L, Taylor C, *et al*. Identifying people with multiple sclerosis in the Canadian primary care sentinel surveillance network. *Mult Scler J Exp Transl Clin* 2019;5:2055217319894360.
30 Marrie RA, Fisk JD, Stadnyk KJ, *et al*. The incidence and prevalence of multiple sclerosis in Nova Scotia, Canada. *Can J Neurol Sci* 2013;40:824–31.
31 Plummer M, Best N, Cowles K, *et al*. CODA: convergence diagnosis and output analysis for MCMC. *R News* 2006;6:7–11.
32 Brooks SP, Gelman A. General methods for monitoring convergence of Iterative simulations. *J Comput Graph Stat* 1998;7:434–55.
33 R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2023.
34 Komárek A, Komárková L. Capabilities of R package mixAK for clustering based on multivariate continuous and discrete longitudinal data. *J Stat Softw* 2014;59:1–38.
35 Murley C, Friberg E, Hillert J, *et al*. Validation of multiple sclerosis diagnoses in the Swedish National Patient Register. *Eur J Epidemiol* 2019;34:1161–9.
36 Al-Sakran LH, Marrie RA, Blackburn DF, *et al*. Establishing the incidence and prevalence of multiple sclerosis in Saskatchewan. *Can J Neurol Sci* 2018;45:295–303.
37 Teljas C, Boström I, Marrie RA, *et al*. Validating the diagnosis of multiple sclerosis using Swedish administrative data in Värmland County. *Acta Neurol Scand* 2021;144:680–6.
38 Marrie RA, O'Mahony J, Maxwell C, *et al*. Incidence and prevalence of MS in children: a population-based study in Ontario, Canada. *Neurology (ECronicon)* 2018;91:e1579–90.
39 Iljicsov A, Milanovich D, Ajtay A, *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.
40 Culpepper WJ, Marrie RA, Langer-Gould A, *et al*. Validation of an algorithm for identifying MS cases in administrative health claims datasets. *Neurology (ECronicon)* 2019;92:e1016–28.
41 Culpepper WJ, Ehrmantraut M, Wallin MT, *et al*. Veterans health administration multiple sclerosis surveillance registry: the problem of case-finding from administrative databases. *J Rehabil Res Dev* 2006;43:17–24.
42 Schmedt N, Khil L, Berger K, *et al*. Incidence of multiple sclerosis in Germany: a cohort study applying different case definitions based on claims data. *Neuroepidemiology* 2017;49:91–8.
43 Lix LM, Walker R, Quan H, *et al*. Features of physician services databases in Canada. *Chronic Dis Inj Can* 2012;32:186–93.

44 Groome PA, McBride ML, Jiang L, *et al*. Lessons learned: it takes a village to understand inter-sectoral care using administrative data across jurisdictions. *Int J Popul Data Sci* 2018;3:440.
45 Canadian Institute for Health Information. Virtual care: a major shift for Canadians receiving physician services. 2022.
46 Canadian Institute for Health Information. *Supply, distribution and migration of physicians in Canada, 2022—data tables*. Ottawa ON: CIHI, 2023.
47 Walton C, King R, Rechtman L, *et al*. Rising prevalence of multiple sclerosis worldwide: insights from the Atlas of MS, third edition. *Mult Scler* 2020;26:1816–21.
48 Danila O, Hirdes JP, Maxwell CJ, *et al*. Prevalence of neurological conditions across the continuum of care based on interRAI assessments. *BMC Health Serv Res* 2014;14:29.
49 Katz A, Enns J, Smith M, *et al*. Population data centre profile: the Manitoba centre for health policy. *Int J Popul Data Sci* 2020;4:1131.
50 Roos LL, Mustard CA, Nicol JP, *et al*. Registries and administrative data: organization and accuracy. *Med Care* 1993;31:201–12.
51 Manitoba Centre for Health Policy. Concept: Manitoba health insurance registry / MCHP research registry - overview [internet]. 2023. Available: http://mchp-appserv.cpe.umanitoba.ca/viewConcept.php?conceptID=1213
52 Marrie RA, Fisk JD, Tremlett H, *et al*. Differences in the burden of psychiatric comorbidity in MS vs the general population. *Neurology (ECronicon)* 2015;85:1972–9.
53 Kroeker K, Widdifield J, Muthukumarana S, *et al*. Model-based methods for case definitions from administrative health data: application to rheumatoid arthritis. *BMJ Open* 2017;7:e016173.
54 Widdifield J, Bernatsky S, Paterson JM, *et al*. Accuracy of canadian health administrative databases in identifying patients with rheumatoid arthritis: a validation study using the medical records of rheumatologists. *Arthritis Care Res (Hoboken)* 2013;65:1582–91.
55 Hutfless S, Jasper RA, Tilak A, *et al*. A systematic review of Crohn's disease case definitions in administrative or claims databases. *Inflamm Bowel Dis* 2023;29:705–15.
56 Lirhus SS, Høivik ML, Moum B, *et al*. Incidence and prevalence of inflammatory bowel disease in Norway and the impact of different case definitions: a nationwide registry study. *Clin Epidemiol* 2021;13:287–94.