

















# Harmonized Data Quality Indicators Maintain Data Quality in Long-Term Safety Studies Using Multiple Sclerosis Registries/Data Sources: Experience from the CLARION Study

Jan Hillert <sup>1</sup>, Helmut Butzkueven <sup>2,3</sup>, Melinda Magyari <sup>4</sup>, Stig Wergeland <sup>5,6</sup>, Nicholas Moore <sup>7</sup>, Merja Soilu-Hänninen <sup>8</sup>, Tjalf Ziemssen <sup>9</sup>, Jens Kuhle <sup>10</sup>, Luigi Pontieri <sup>11</sup>, Lars Forsberg <sup>1</sup>, Jan Harald Aarseth <sup>5</sup>, Chao Zhu <sup>2</sup>, Nicholas Sicignano <sup>12</sup>, Vasili Mushnikov <sup>13</sup>, Irene Bezemer <sup>14</sup>, Meritxell Sabidó <sup>15</sup>

<sup>1</sup>Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden; <sup>2</sup>Department of Neuroscience, Central Clinical School, Monash University, Melbourne, Victoria, Australia; <sup>3</sup>The Alfred Hospital, Melbourne, Victoria, Australia; <sup>4</sup>Danish Multiple Sclerosis Center and Danish Multiple Sclerosis Registry, Department of Neurology, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark; <sup>5</sup>Norwegian MS Registry and Biobank, Department of Neurology, Haukeland University Hospital, Bergen, Norway; <sup>6</sup>Department of Clinical Medicine, University of Bergen, Bergen, Norway; <sup>7</sup>Bordeaux PharmacoEpi (BPE), Université de Bordeaux, Bordeaux, France; <sup>8</sup>Turku University Hospital Neurocenter and Division of Clinical Neurosciences, University of Turku, Turku, Finland; <sup>9</sup>Center of Clinical Neuroscience, Department of Neurology, University Clinic Carl Gustav Carus, Dresden University of Technology, Dresden, Germany; <sup>10</sup>MS Centre and Research Center for Clinical Neuroimmunology and Neuroscience Basel (RC2NB), Departments of Neurology, Biomedicine and Clinical Research, University Hospital Basel, University of Basel, Basel, Switzerland; <sup>11</sup>Danish Multiple Sclerosis Registry, Department of Neurology, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark; <sup>12</sup>Health ResearchTx LLC, Trevose, PA, USA; <sup>13</sup>QVIA, Helsinki, Finland; <sup>14</sup>QVIA, Amsterdam, the Netherlands; <sup>15</sup>Merck Healthcare KGaA, Darmstadt, Germany

Correspondence: Meritxell Sabidó, Merck Healthcare KGaA, Frankfurter Str. 250, Darmstadt, 64293, Germany, Email [meritxell.sabido-espin@merckgroup.com](mailto:meritxell.sabido-espin@merckgroup.com)

**Purpose:** Understanding the long-term safety of disease-modifying therapies for multiple sclerosis (MS) in routine clinical practice can be undertaken through registry-based studies. However, variability of data quality across such sources poses the challenge of data fit for regulatory decision-making. CLARION, a non-interventional cohort safety study of cladribine tablets, combines aggregated data from MS registries/data sources, except in Germany (which utilizes primary data collection). We describe the application of key data quality indicators (DQIs) within CLARION to evaluate data quality over time, as recommended by the European Medicines Agency (EMA) guideline on registry-based studies.

**Methods:** DQIs were defined with participating registries/sources; they were used to assess data quality according to the EMA Data Quality Framework, addressing consistency, accuracy, completeness, and study representativeness. DQIs were associated with potential remedial measures if data quality was not met. DQIs were summarized overall and for individual MS registries/data sources to November 1, 2022.

**Results:** A total of 28 DQIs were analyzed using data from 5069 patients arising from eight MS registries/data sources and 14 countries. The Representativeness DQIs showed that 72.0% of patients were female, median age at MS diagnosis was 29.0 to 43.3 years, and 93.5% had relapsing-remitting MS. Consistency DQIs showed a total of 2899 patients had achieved at least two years of follow-up; 6.9% did not have any recorded visits during this timeframe. Discrepant values were assessed as part of Accuracy DQIs, and improvements over time were noted for recorded dates of MS onset and diagnosis. Regarding Completeness DQIs, 191/5069 (3.8%) patients were lost to follow-up.

**Conclusion:** The application of 28 DQIs within the CLARION study has helped with understanding, not only intrinsic and question-specific determinants of data quality, but also tracking the quality of post-authorization safety data obtained from MS registries/data sources, thereby providing a foundation for the regulatory decision-making process.

**Keywords:** cladribine tablets, multiple sclerosis, safety, fingolimod

## Introduction

The characterization of the long-term safety of treatments in routine clinical practice, as part of post-approval commitments to regulatory authorities, can be enhanced by the collection and analysis of drug utilization and safety data from disease registries and other real-world data sources. Numerous registries and data sources are well established in the multiple sclerosis (MS) setting, a potentially disabling disease requiring life-long treatment. A key strength of MS registries/data sources is that, with long-term continuous monitoring and minimal loss to follow-up, they amass a large quantity of high-quality data across several sites.<sup>1</sup> In parallel, several MS registries have started to collect data using a unified approach that allows for harmonization across different sources,<sup>2,3</sup> thereby supporting the efforts of researchers from third parties to conduct aggregated data analysis.

The CLARION study, the design of which is published elsewhere,<sup>4</sup> was initiated to assess the long-term safety of cladribine tablets. Since first approval in 2017, an estimated 101,132 people have received cladribine tablets for the treatment of MS with 251,900 cumulative years of exposure (as of end-June 2024; Merck, data on file). In brief, CLARION is an ongoing, multi-country, comparative, non-interventional cohort study that includes patients newly initiating cladribine tablets or fingolimod (main comparator) for relapsing-remitting MS (target N = 4000 per group). The study uses a study-specific common data model with local analysis and common programs to transform and combine real-world data from multiple sources, including MS registries, medical claims and, in Germany, primary data collection. Results of the first pre-planned interim analysis (cut-off date of April 1, 2020) have recently been published.<sup>5</sup> At the time of study set-up, the European Medicines Agency (EMA) Patient Registries Initiative hosted a workshop on MS registries.<sup>6</sup> One of the workshop recommendations, in terms of data quality for registry-based studies such as CLARION, was the definition of key data quality indicators (DQIs). The same recommendation is associated with remedial measures if acceptable levels of quality are not met and can be found in the EMA guideline on registry-based studies.<sup>7</sup> Furthermore, the Data Quality Framework (DQF) for European Union medicines regulation<sup>8</sup> provides detail on data quality metrics that can be applied to intrinsic determinants, which pertain to aspects that are inherent to a specific dataset, and question-specific determinants, which pertain to aspects of data quality that cannot be defined independently to a specific question, to derive assessments of one or more categories of data quality. Different dimensions address distinct data quality questions, and the sum of the independent features of the dimensions reveals the data quality of the data sources. The EMA's dimensions are reliability, extensiveness, coherence timeliness, and relevance, which, for this study, refers to the suitability of registries to answer the objectives posed in CLARION. The EMA DQF also considers representativeness as a critical element for maximizing the use of real-world data in regulatory decision-making.

The CLARION study adopted the EMA recommendations and defined a set of DQIs to assess some of these data quality dimensions as part of the study, thereby maintaining data that are fit for regulatory decision-making on the safety of cladribine tablets. This method also offers a unique insight into the overall data quality generated by participating MS registries/data sources, in absolute terms and relative to each other, and consequently the ability to combine these data into one study.

## Methods

### Development of DQIs

DQIs were defined in collaboration with participating MS registries/data sources, the study principal investigators, and experts from the marketing authorization holder. The DQIs were tested in a subset of registries and progressively scaled up to other participating registries/data sources when they joined the CLARION study. Firstly, the four main categories to be assessed were agreed as representativeness, consistency, accuracy, and completeness. Secondly, specific DQIs were defined for each category. The EMA DQF includes foundational determinants previously evaluated through a detailed questionnaire that included both scientific and operational aspects during the CLARION start-up activities. Timeliness, one of the EMA DQF (five) dimensions of data quality, was evaluated at the study feasibility stage before approval of the study protocol and was further confirmed during the questionnaire. Traceability, a feature of reliability, was addressed as part of the data management plan.

For each MS registry/data source, remedial actions were mapped to understand what measures could be triggered if data quality was not met (eg, data verification, personnel training, and center/hospital feedback). DQIs were analyzed when the MS registry/data source joined the study and yearly afterwards. The DQIs were established to assess data quality produced by participating MS registries/data sources. If a change in the DQI within a registry was observed, the research team interacted with the MS registry/data source to understand potential quality concerns to determine if the source data or study definitions needed improvement and, when applicable, remedial actions were activated.

## MS Registries/Data Sources

The present report concerns data from patients included to CLARION from eight participating MS registries/data sources: MS Documentation System 3D (MSDS3D; Germany), which utilizes primary data collection;<sup>9</sup> and the Danish MS Registry, Finnish MS Registry, MS Database (MSBase, multiple countries), Norwegian MS Registry and Biobank (NMSRB), Swedish MS Registry, Swiss MS Cohort, and US Department of Defense (DoD), all of which utilize secondary data collection. All transfer aggregated data to the research team for purposes of analysis (except MSDS3D and the Swiss MS Cohort, both of which transfer patient-level data).

## Data Quality Indicators

A total of 28 DQIs were grouped to describe the four categories of data quality:

- **Representativeness** (5 DQIs): [*dimension of relevance*] To assess the distribution and representativeness of key patient characteristics within the study population in each MS registry/data source.
- **Consistency** (4 DQIs): [*dimension of coherence*] To evaluate uniformity of core data elements entered over time by considering data-recording density and frequency over time.
- **Accuracy** (7 DQIs): [*dimension of reliability*] To assess how well the data are entered by identifying potentially discrepant values; a value is considered discrepant when there is an issue, such as an overlap, with the chronological order of dates. Duplicate values are also identified.
- **Completeness** (12 DQIs): [*dimension of extensiveness*] To assess how much data are missing by describing the proportion of missing values in the key outcomes and variables.

The derivation details for the DQIs ([Supplementary Table 1](#)), and the imputation algorithm for any partial MS treatment start dates, are shown in [Supplementary Materials](#).

## Analysis

At each data cut-off from May 1, 2019 onwards, the DQIs are summarized overall and for individual MS registries/data sources as part of regular Data Quality Assessment Reports (DQARs); data at the time of writing are available to the most recent report concerning a cut-off date of November 1, 2022.

## Results

### Population

A total of 28 DQIs covering Representativeness, Consistency, Accuracy, and Completeness were analyzed using data from 5,069 patients with confirmed eligibility (2,958 [58.4%] patients in the cladribine cohort and 2,111 [41.6%] patients in the fingolimod cohort) arising from eight MS registries/data sources and 14 countries, using November 1, 2022, as the data cut-off date ([Table 1](#)).

### Representativeness

The results for the five Representativeness DQIs are presented in [Tables 2–4](#).

**Table 1** Number of Patients in the CLARION Study Population with Confirmed Eligibility

MS Registry/Data Source	Study Population (Confirmed Eligibility), N (%)		
	Cladribine Cohort (N=2958)	Fingolimod Cohort (N=2111)	Total (N=5069)
Danish MS Registry	255 (8.6)	500 (23.7)	755 (14.9)
Finnish MS Registry	229 (7.7)	194 (9.2)	423 (8.3)
MSBase (multiple countries)	640 (21.6)	723 (34.2)	1,363 (26.9)
MSDS3D (Germany)	617 (20.9)	221 (10.5)	838 (16.5)
Norwegian MS Registry and Biobank	735 (24.8)	210 (9.9)	945 (18.6)
Swedish MS Registry	261 (8.8)	77 (3.6)	338 (6.7)
Swiss MS Cohort	52 (1.8)	39 (1.8)	91 (1.8)
US Department of Defense	169 (5.7)	147 (7.0)	316 (6.2)

**Notes:** Cut-off date for most recent data quality indicators: November 1, 2022.

**Abbreviations:** MS, multiple sclerosis; MSBase, MS Database; MSDS3D, MS Management System 3D.

Sex distribution, presented as the percentage of female patients (Representativeness DQI 1), was similar in the cladribine and fingolimod cohorts but ranged across MS registries/data sources from 61.5% in the Swiss MS cohort to 80.1% in the US DoD data source; overall, 72.0% of patients were female.

Median age at MS onset, as shown in Representativeness DQI 2, ranged from 28.0 to 32.0 years over seven contributing MS registries/data sources (no data concerning age at MS onset were available from the US DoD data source). Across the cladribine and fingolimod cohorts, the median age at MS onset ranged from 27.0 to 32.0 years and 29.0 to 35.0 years, respectively. The median age at MS diagnosis (Representativeness DQI 3) ranged from 29.0 to 43.3 years over all eight MS registries/data sources. Across the cladribine and fingolimod cohorts, the median age at MS diagnosis ranged from 28.0 to 49.3 years and from 30.5 to 38.4 years, respectively, with the study population from the US DoD data source tending to be slightly older.

Relapsing-remitting MS was the predominant (93.5%) clinical course in all data sources (Representativeness DQI 4).

Details regarding prior disease-modifying therapies (DMTs) are illustrated in Representativeness DQI 5; overall, 30.0% of patients were treatment naïve (including 35.6% and 22.2% of patients in the cladribine and fingolimod cohorts, respectively). Considerable variability across MS registries/data sources was noted.

## Consistency

The results for the Consistency DQIs are presented in Table 5.

A requirement of the Consistency DQIs was such that at least one year of follow-up per patient was available (measured by Consistency DQI 1), which was observed for 4,137 patients at the current data cut-off. Within this requirement, 20.8% of patients ( $n/N = 860/4,137$ ) had no recorded visit, and some differences were noted between individual MS registries/data sources (Figure 1). Two years of follow-up per patient was a requirement of Consistency DQI 2. A total of 2,899 patients achieved this, but 6.9% did not have any recorded visits during the last two years of follow-up.

Overall, 24.0% of 2,927 patients (excluding patients from NMSRB and the Finnish MS Registry because visit dates are not recorded) with at least one year of follow-up had less than one visit per year, on average, during follow-up (Consistency DQI 3). Considerable variability across MS registries/data sources was noted. For Consistency DQI 4, the average number of recorded lymphocyte count measurements per patient per year in the cladribine cohort was 1.0. However, not all MS registries/data sources routinely record lymphocyte counts.

**Table 2** Representativeness DQIs: Patient Demographics and Characteristics in CLARION

MS Registry/Data Source	DQI 1; Female Sex (%)			DQI 2; Median (IQR) Age at MS Onset			DQI 3; Median (IQR) Age at MS Diagnosis		
	Cladribine Cohort (N=2958)	Fingolimod Cohort (N=2111)	Total (N=5069)	Cladribine Cohort (N=2789)	Fingolimod Cohort (N=1964)	Total (N=4753)	Cladribine Cohort (N=2958)	Fingolimod Cohort (N=2111)	Total (N=5069)
Danish MS Registry	66.3	67.6	67.2	30.0 (24.0, 37.0)	30.0 (23.0, 37.0)	30.0 (24.0, 37.0)	32.0 (27.0, 40.0)	33.0 (25.0, 41.0)	33.0 (26.0, 41.0)
Finnish MS Registry	79.5	78.4	79.0	27.0 (22.0, 34.0)	29.0 (22.0, 35.0)	28.0 (22.0, 34.0)	28.0 (24.0, 36.0)	30.5 (24.0, 37.0)	29.0 (24.0, 37.0)
MSBase (multiple countries)	73.3	70.8	72.0	31.0 (25.0, 40.0)	29.0 (22.0, 35.0)	30.0 (23.0, 37.0)	34.0 (28.0, 43.0)	31.0 (24.0, 38.0)	33.0 (26.0, 40.0)
MSDS3D (Germany)	77.6	65.6	74.5	28.0 (23.0, 36.0)	30.0 (23.0, 37.0)	28.0 (23.0, 36.0)	30.0 (24.0, 37.0)	31.0 (24.0, 38.8)	30.0 (24.0, 37.0)
Norwegian MS Registry and Biobank	69.4	67.1	68.9	32.0 (26.0, 40.0)	33.0 (27.0, 40.0)	32.0 (26.0, 40.0)	35.0 (28.0, 43.0)	36.0 (29.0, 43.0)	35.0 (28.0, 43.0)
Swedish MS Registry	75.5	62.3	72.5	28.0 (23.0, 33.0)	29.0 (24.0, 34.0)	28.0 (23.0, 34.0)	30.0 (25.0, 36.0)	31.0 (27.0, 37.0)	30.0 (26.0, 36.0)
Swiss MS Cohort	61.5	61.5	61.5	30.0 (26.0, 39.5)	35.0 (27.0, 41.8)	31.0 (26.0, 41.0)	31.0 (26.0, 41.5)	38.0 (27.0, 48.5)	35.5 (27.0, 43.0)
US Department of Defense	82.2	77.6	80.1	NA			49.3 (37.8, 57.6)	38.4 (30.4, 51.4)	43.3 (33.8, 54.9)

**Notes:** Cut-off date for most recent data quality indicators: November 1, 2022.

**Abbreviations:** DQI, data quality indicator; IQR, interquartile range; MS, multiple sclerosis; MSBase, MS Database; MSDS3D, MS Management System 3D; NA, not applicable.

**Table 3** Representativeness DQI 4: MS Clinical Course at Inclusion in CLARION

MS Registry/ Data Source	DQI 4; MS Clinical Course at Inclusion (% of Population)												
	RRMS			SPMS			PPMS			PRMS			Unknown
	Cladribine Cohort (N=2789)	Fingolimod Cohort (N=1964)	Total (N=4753)	Cladribine Cohort (N=2054)	Fingolimod Cohort (N=1754)	Total (N=3808)	Cladribine Cohort (N=2789)	Fingolimod Cohort (N=1964)	Total (N=4753)	Cladribine Cohort (N=2054)	Fingolimod Cohort (N=1754)	Total (N=3808)	Total (N=4753)
Danish MS Registry <sup>a</sup>	97.6	91.2	93.4	1.6	8.8	6.4	0.8	0.0	0.3	NA			0.0
Finnish MS Registry	96.9	94.3	95.7	2.6	5.7	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
MSBase (multiple countries)	84.1	93.4	89.0	6.1	2.4	4.1	1.1	0.3	0.7	0.9	0.1	0.5	5.7
MSDS3D (Germany)	97.2	100	98.0	2.8	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Norwegian MS Registry and Biobank <sup>b</sup>	94.7	94.8	94.7	NA			1.0	0.0	0.7	NA			4.6
Swedish MS Registry <sup>a</sup>	94.6	94.8	94.7	4.2	3.9	4.1	1.1	1.3	1.2	NA			0.0
Swiss MS Cohort <sup>a</sup>	86.5	100	92.3	11.5	0.0	6.6	1.9	0.0	1.1	NA			0.0
US Department of Defense <sup>c</sup>	NA												

**Notes:** Cut-off date for most recent data quality indicators: November 1, 2022. <sup>a</sup>Danish and Swedish MS Registries, and the Swiss MS Cohort, do not have PRMS among possible MS course choices. <sup>b</sup>Norwegian MS Registry and Biobank did not report data on SPMS or PRMS. <sup>c</sup>US Department of Defense did not report data concerning DQI 4.

**Abbreviations:** DMT, disease-modifying therapy; DQI, data quality indicator; MS, multiple sclerosis; MSBase, MS Database; MSDS3D, MS Management System 3D; NA, not applicable; PPMS, primary progressive MS; PRMS, primary relapsing MS; RRMS, relapsing-remitting MS; SPMS, secondary progressive MS.

**Table 4** Representativeness DQI 5: Number of Previous DMTs in CLARION

MS Registry/Data Source	DQI 5; Number of Previous DMTs <sup>a</sup> (% of Population)											
	Cladribine Cohort (N=2958)				Fingolimod Cohort (N=2111)				Total (N=5069)			
	0	1	2	≥3	0	1	2	≥3	0	1	2	≥3
Danish MS Registry	18.0	27.8	24.3	29.8	16.4	36.6	25.2	21.8	17.0	33.6	24.9	24.5
Finnish MS Registry	35.8	31.0	18.8	14.4	18.6	38.7	24.2	18.6	27.9	34.5	21.3	16.3
MSBase (multiple countries)	34.8	38.1	17.7	9.4	18.9	65.0	13.4	2.6	26.4	52.4	15.4	5.8
MSDS3D (Germany)	25.3	37.1	19.3	18.3	19.0	43.4	20.8	16.7	23.6	38.8	19.7	17.9
Norwegian MS Registry and Biobank	49.3	28.3	13.6	8.8	26.7	36.2	20.0	17.1	44.2	30.1	15.0	10.7
Swedish MS Registry	41.0	20.7	16.5	21.8	24.7	33.8	20.8	20.8	37.3	23.7	17.5	21.6
Swiss MS Cohort	36.5	25.0	11.5	26.9	56.4	25.6	12.8	5.1	45.1	25.3	12.1	17.6
US Department of Defense	34.9	37.3	17.8	10.1	51.0	23.8	15.6	9.5	42.4	31.0	16.8	9.8

**Notes:** Cut-off date for most recent data quality indicators: November 1, 2022. <sup>a</sup>Per protocol, patients were excluded from this study if they had received fingolimod prior to initiating cladribine tablets or received cladribine tablets prior to initiating fingolimod.

**Abbreviations:** DMT, disease-modifying therapy; DQI, data quality indicator; MS, multiple sclerosis; MSBase, MS Database; MSDS3D, MS Management System 3D.

**Table 5** Consistency DQIs in CLARION

MS Registry/Data Source	DQI 1; No Visits During Past 1 Year—Total Population (N=4137) (%)	DQI 2; No Visits During Past 2 Years—Total Population (N=2899) (%)	DQI 3; Less Than 1 Visit/Year During Follow-Up—Total Population <sup>a</sup> (N=2927) (%)	DQI 4; Average Number of Recorded Lymphocyte Count Measurements/Patient/Year—Cladribine Cohort <sup>b</sup> (N=1523)						
				Mean (SD)	None	0-1	1-2	2-3	3-4	>4
Danish MS Registry	16.6	3.5	33.4	NA						
Finnish MS Registry	6.6	2.1	NA	2.7 (1.7)	2.7	7.0	24.2	33.9	21.5	10.8
MSBase (multiple countries)	26.3	12.4	28.4	1.1 (1.7)	42.4	29.2	9.5	5.2	6.6	7.2
MSDS3D (Germany)	6.8	0.0	4.5	0.3 (0.6)	74.0	18.7	4.8	1.6	0.7	0.2
Norwegian MS Registry and Biobank	43.4	11.8	NA	NA						
Swedish MS Registry	6.5	3.0	43.1	1.2 (1.0)	29.3	18.4	32.8	15.5	2.3	1.7
Swiss MS Cohort	8.8	4.7	48.8	1.5 (1.9)	44.2	32.6	11.6	4.7	4.7	2.3
US Department of Defense	7.0	2.8	0.0	1.1 (3.0)	79.9 <sup>c</sup>	4.1	1.8	3.0	1.2	10.1

**Notes:** Cut-off date for most recent data quality indicators: November 1, 2022. <sup>a</sup>Finnish MS Registry and Norwegian MS Registry and Biobank did not contribute to DQI 3.

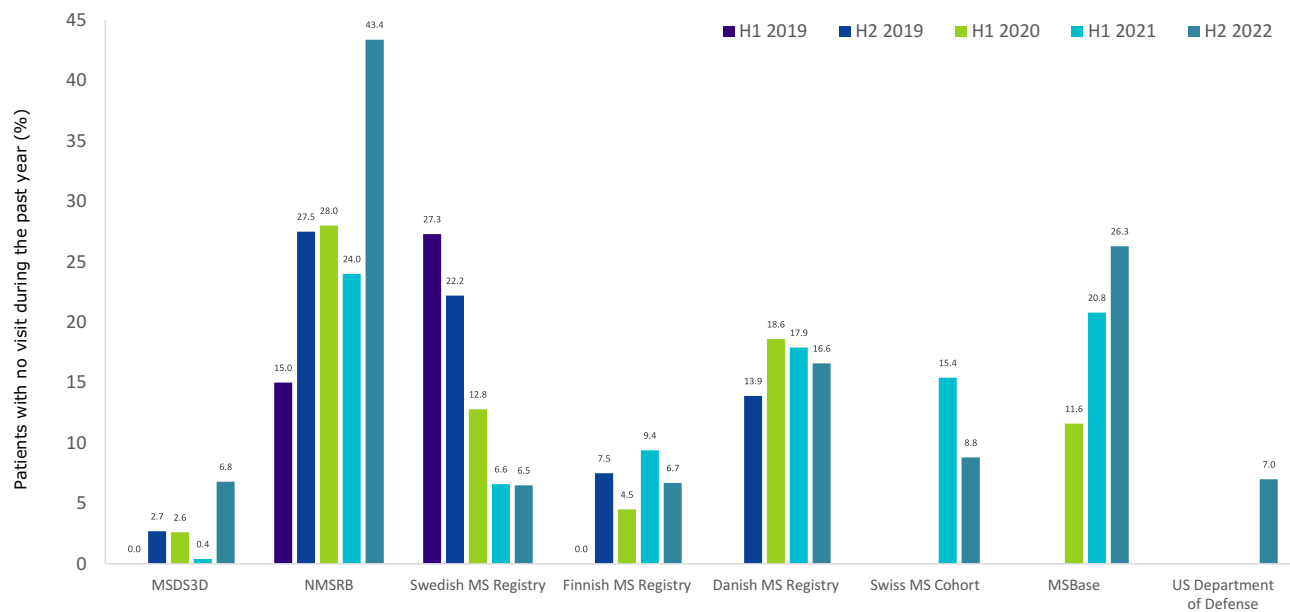
<sup>b</sup>Danish MS Registry and Norwegian MS Registry and Biobank did not contribute to DQI 4 as these data sources did not collect data on lymphocyte counts. <sup>c</sup>US Department of Defense did not include patients receiving care in the civilian sector—only those within US Department of Defense treatment facilities (~30% of the source population)—so the results can only be partially evaluated for DQI 4.

**Abbreviations:** DQI, data quality indicator; MS, multiple sclerosis; MSBase, MS Database; MSDS3D, MS Management System 3D; NA, not applicable; SD, standard deviation.

## Accuracy

The results, showing the accuracy of data collection in CLARION, are presented in [Table 6](#).

Discrepant MS onset dates for Accuracy DQI 1 were detected for 1.1% of 4,635 patients at the current data cut-off, while discrepant MS diagnosis dates (Accuracy DQI 2) were detected for 0.6% of 4,661 evaluable patients (not including



**Figure 1** Patients with no visit during the past year at data cut-off (Consistency DQI 1).

**Notes:** H1 2019, May 1, 2019, data cut-off; H2 2019, November 1, 2019, data cut-off; H1 2020, May 1, 2020, data cut-off; H1 2021, May 1, 2021, data cut-off; H2 2022, November 1, 2022, data cut-off.

**Abbreviations:** DQI, data quality indicator; MS, multiple sclerosis; MSBase, MS Database; MSDS3D, Multiple Sclerosis Management System 3D; NMSRB, Norwegian Multiple Sclerosis Registry and Biobank.

the US DoD data source; such data were not evaluable as the date of MS onset was not available and the reported date of MS diagnosis concerns the earliest date when a MS diagnosis code was recorded rather than the exact date of diagnosis). Both DQIs 1 and 2 showed improvement over time. Specifically, the percentage of patients with data discrepancies decreased over the study duration (Figure 2).

Overall, 1.5% of 5,069 patients had records for MS treatment stop dates at the current data cut-off that were considered discrepant (Accuracy DQI 3). The majority of these were reported for the US DoD data source (15.2%);

**Table 6** Accuracy DQIs in CLARION

MS Registry/Data Source	DQI 1; Discrepant MS Onset Date (N=4635) (%) <sup>a</sup>	DQI 2; Discrepant MS Diagnosis Date (N=4661) (%) <sup>a</sup>	DQI 3; Discrepant MS Treatment Stop Dates (N=5069) (%)	DQI 4; Discrepant Treatment Stop Dates (Cladribine Cohort) (N=359) (%)	DQI 5; Discrepant Treatment Stop Dates (Fingolimod Cohort) (N=684) (%)	DQI 6; Discrepant Stop Date of AESI (N=29) (%) <sup>b</sup>	DQI 7; Duplicated MS Treatment Recordings (N=5069) (%)
Danish MS Registry	1.9	0.0	2.0	12.5	3.0	NA	0.0
Finnish MS Registry	3.4	2.6	0.0	15.0	3.0	NA	5.0
MSBase (multiple countries)	0.6	0.0	0.2	0.0	2.8	0.0	0.0
MSDS3D (Germany)	0.0	0.0	0.1	0.0	0.0	5.9	0.2
Norwegian MS Registry and Biobank	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Swedish MS Registry	5.2	4.8	2.7	26.7	0.0	NA	0.0

(Continued)



**Table 6** (Continued).

MS Registry/Data Source	DQI 1; Discrepant MS Onset Date (N=4635) (%) <sup>a</sup>	DQI 2; Discrepant MS Diagnosis Date (N=4661) (%) <sup>a</sup>	DQI 3; Discrepant MS Treatment Stop Dates (N=5069) (%)	DQI 4; Discrepant Treatment Stop Dates (Cladribine Cohort) (N=359) (%)	DQI 5; Discrepant Treatment Stop Dates (Fingolimod Cohort) (N=684) (%)	DQI 6; Discrepant Stop Date of AESI (N=29) (%) <sup>b</sup>	DQI 7; Duplicated MS Treatment Recordings (N=5069) (%)
Swiss MS Cohort	0.0	0.0	1.1	0.0	10.0	0.0	0.0
US Department of Defense	NA	NA	15.2 <sup>c</sup>	14.8 <sup>c</sup>	19.5 <sup>c</sup>	NA	2.5

**Notes:** Cut-off date for most recent data quality indicators: November 1, 2022. <sup>a</sup>DQI 1 and DQI 2 were not evaluable for the US Department of Defense because MS onset date is not available in the data source and the available MS diagnosis date is not the exact MS diagnosis date but the earliest date when an MS diagnosis code was recorded. <sup>b</sup>DQI 6 was not evaluated for the Danish, Finnish, and Swedish MS Registries, and for the US Department of Defense. <sup>c</sup>DQI 3, DQI 4, and DQI 5 use a proxy definition for stop date equal to the start date plus number of days of supply, which can result in an overlap of MS treatments.

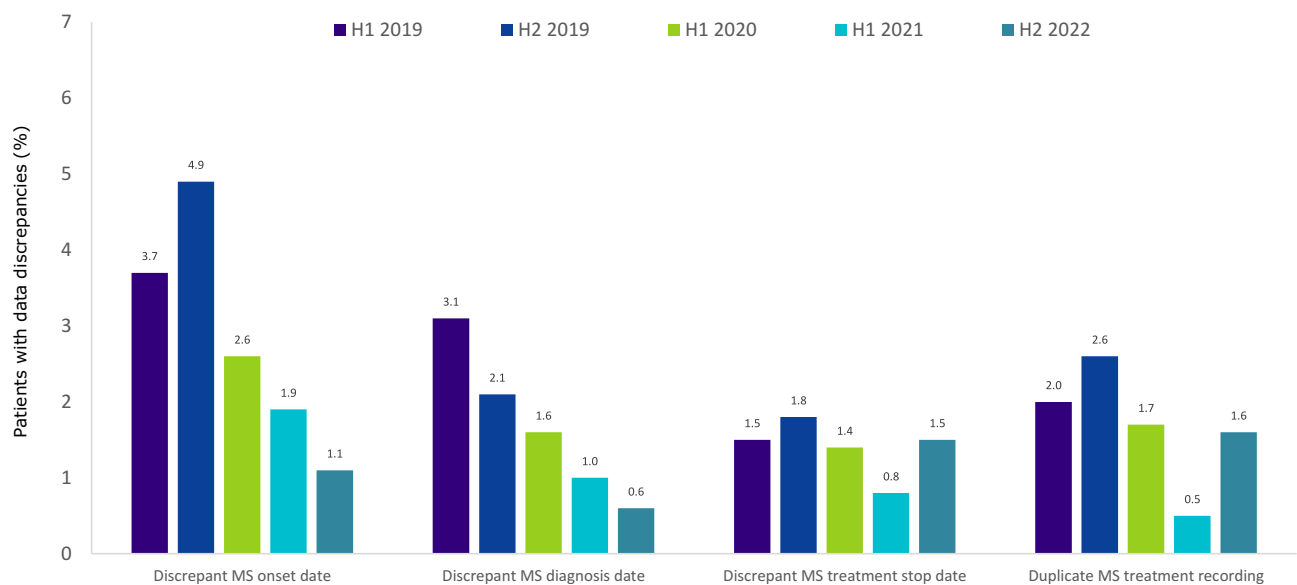
**Abbreviations:** AESI, adverse events of special interest; DQI, data quality indicator; MS, multiple sclerosis; MSBase, MS Database; MSDS3D, MS Management System 3D; NA, not applicable.

however, this is a claims-based database, and a proxy definition was used to define the stop date (which was equal to the start date plus number of days of supply). In the cladribine cohort, 10.0% of patients who switched to another DMT were classified as having discrepant study treatment stop dates; this figure was 5.3% in the fingolimod cohort. Accuracy DQI 6 considered discrepant stop dates of adverse events of special interest (AESI), but only limited data were available. Of AESIs with recorded stop dates, a total of 29 AESIs were reported; only one had a stop date considered discrepant.

In the total study population, 1.6% of 5,069 patients had at least one instance of MS treatment recordings that were considered duplicates at the current data cut-off (Accuracy DQI 7). These were observed in three data sources: Germany's MSDS3D (0.2% of patients had duplicates), the US DoD data source (2.5%), and the Finnish MS Registry (5.0%).

## Completeness

The results for Completeness DQIs are presented in Table 7 and Figure 3. Several data sources did not contribute DQI information for specific DQIs; the findings are therefore reported from data sources that did contribute data.

**Figure 2** Patients with data discrepancies at data cut-off (selected Accuracy DQIs).

**Notes:** H1 2019, May 1, 2019, data cut-off; H2 2019, November 1, 2019, data cut-off; H1 2020, May 1, 2020, data cut-off; H1 2021, May 1, 2021, data cut-off; H2 2022, November 1, 2022, data cut-off.

**Abbreviations:** DQI, data quality indicator; MS, multiple sclerosis.

**Table 7** Completeness DQIs in CLARION

DQI	Danish MS Registry	Finnish MS Registry	MSBase (multiple countries)	MSDS3D (Germany)	Norwegian MS Registry and Biobank	Swedish MS Registry	Swiss MS Cohort	US Department of Defense
DQI 1; records of MS treatment with missing start date (%)								
Only day missing	0.0	0.0	0.0	14.4	0.0	0.0	1.0	0.0
Both month and day missing	0.0	0.0	0.0	4.9	0.0	0.0	0.0	0.0
Completely missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Unknown	NA	NA	NA	2.1	NA	NA	NA	NA
DQI 2; records of AESI with missing details (%) <sup>a</sup>								
	NA	NA	NA	43.9	NA	NA	NA	NA
DQI 3; records of lymphocyte count with missing date (cladribine cohort) (%) <sup>b</sup>								
Only day missing	NA	0.0	0.0	0.0	NA	0.0	0.0	0.0
Both month and day missing	NA	0.0	0.0	0.0	NA	0.0	0.0	0.0
Completely missing	NA	0.0	21.2	0.0	NA	0.0	0.0	0.0
DQI 4; patient dropout by year (%)								
2017	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2018	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2019	0.3	0.2	1.0	0.6	0.0	0.0	0.0	0.3
2020	0.3	0.0	5.4	1.4	0.0	0.0	2.2	0.9
2021	0.3	0.0	0.1	3.2	0.2	0.0	0.0	0.6
2022	0.1	0.0	0.1	4.4	0.1	0.3	1.1	0.0
DQI 5; records of severe lymphopenia with missing stop date (%) <sup>c</sup>								
	NA	NA	100	1.0	0.0	NA	NA	NA
DQI 6; records of MS treatment with missing stop date (%) <sup>d</sup>								
Only day missing	NA	14.5	0.0	16.8	0.0	NA	14.1	NA
Both month and day missing	NA	0.8	0.0	3.8	0.0	NA	2.4	NA
Completely missing	NA	0.0	0.9	0.0	0.0	NA	0.8	NA
Unknown	NA	NA	NA	2.1	NA	NA	NA	NA
DQI 7; records of missing reason for treatment discontinuation (cladribine cohort) (%) <sup>e</sup>								
	0.0	28.1	21.2	0.6	0.0	0.0	0.0	NA
DQI 8; records of missing reason for treatment discontinuation (fingolimod cohort) (%) <sup>e</sup>								
	0.0	0.0	19.8	0.0	0.0	0.0	0.0	NA

(Continued)

Table 7 (Continued).

DQI	Danish MS Registry	Finnish MS Registry	MSBase (multiple countries)	MSDS3D (Germany)	Norwegian MS Registry and Biobank	Swedish MS Registry	Swiss MS Cohort	US Department of Defense
DQI 9; patients with missing birth date (%)								
Only day missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Both month and day missing	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0
Completely missing	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0
DQI 10; patients with missing sex (%)								
Only day missing	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DQI 11; patients with missing MS onset date (%) <sup>f</sup>								
Only day missing	0.3	21.3	0.0	41.6	76.9	18.3	26.4	NA
Both month and day missing	39.2	10.6	0.0	27.8	23.1	15.4	22.0	NA
Completely missing	0.0	1.9	2.1	0.0	0.0	2.4	2.1	NA
Unknown	NA	NA	NA	8.5	NA	NA	NA	NA
DQI 12; patients with missing MS diagnosis date (%) <sup>f</sup>								
Only day missing	0.3	6.4	0.0	44.0	0.0	2.1	17.6	NA
Both month and day missing	26.0	3.3	0.0	19.6	0.0	3.0	13.2	NA
Completely missing	0.0	0.0	4.9	0.0	0.0	1.8	1.1	NA
Unknown	NA	NA	NA	2.0	NA	NA	NA	NA

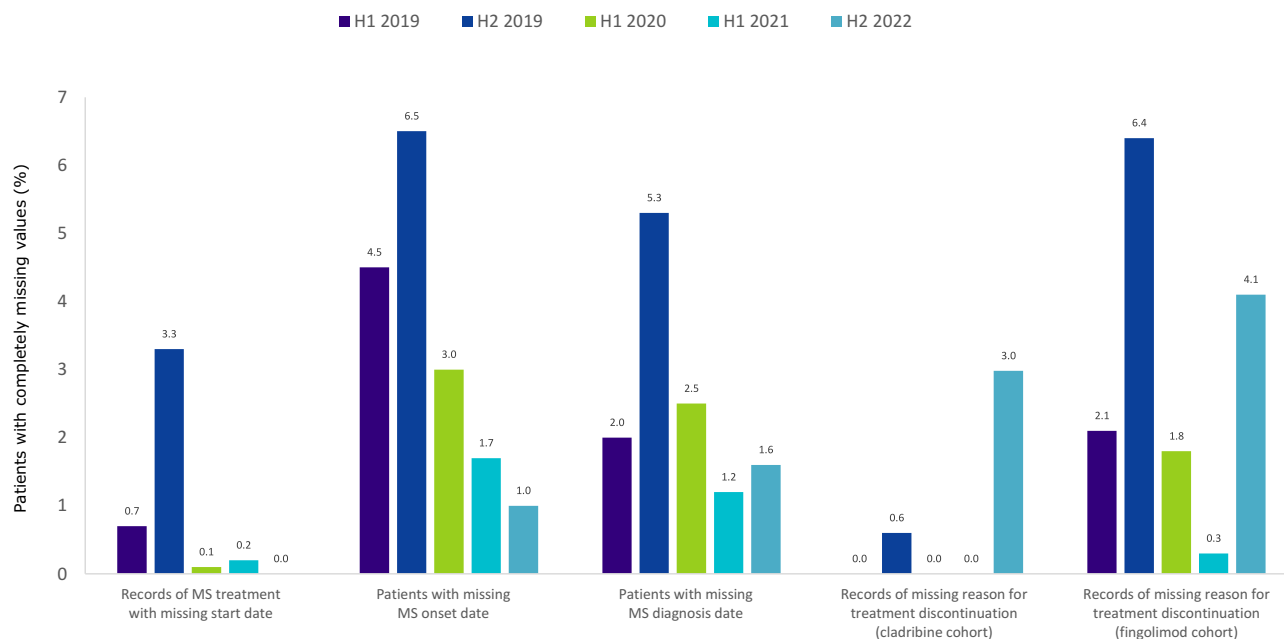
**Notes:** Cut-off date for most recent data quality indicators: November 1, 2022. <sup>a</sup>Evaluated for MSDS3D only (the only data source that collected relevant information). <sup>b</sup>Not evaluated for the Danish MS Registry or the Norwegian MS Registry and Biobank as these data sources did not record lymphocyte counts. Evaluable for a subset of patients with laboratory values (~30%) in the US Department of Defense. <sup>c</sup>Not evaluated in the Danish, Finnish, and Swedish MS registries as information on severe lymphopenia was not available. Data from the US Department of Defense were not available at the time of analysis. It was not evaluable in the Swiss MS Cohort as no lymphopenia episodes were included in the denominator. <sup>d</sup>Not evaluated in the Danish and Swedish MS registries, or the US Department of Defense. <sup>e</sup>Not reported for the US Department of Defense as the data source did not record treatment discontinuations. <sup>f</sup>Not reported for the US Department of Defense as the data source did not record MS onset date and the exact date of MS diagnosis.

**Abbreviations:** AESI, adverse events of special interest; DQI, data quality indicator; MS, multiple sclerosis; MSBase, MS Database; MSDS3D, MS Management System 3D; NA, not applicable.

In Completeness DQI 1, there were no missing MS treatment start dates for the total of 9,470 MS treatment episodes at the current data cut-off. Incomplete AESI details (Completeness DQI 2), which could only be evaluated in MSDS3D, were classified in 18 AESIs out of 41 reports (43.9%). Among MS registries/data sources that collected data on severe lymphopenia, as part of Completeness DQI 3, an evaluation of missing lymphocyte count measurement dates in the cladribine cohort revealed that only MSBase had completely missing dates of lymphocyte count measurements in some instances (21.2% [484/2,278]).

At the current data cut-off, 191/5,069 (3.8%) patients were classified as having been lost to follow-up over the years 2019–2022 inclusive, as shown in Completeness DQI 4. Completeness DQI 5, concerning missing severe lymphopenia stop dates, was evaluated for MSDS3D and MSBase. Missing values were detected in five of 123 episodes, with one case in MSDS3D and four cases in MSBase.

MS treatment stop dates (Completeness DQI 6) were completely missing in a few cases (0.1% of 7,128 MS treatment episodes evaluated). Reasons for cladribine treatment discontinuation were not recorded for 3.0% of 2,719 discontinued treatment episodes (Completeness DQI 7). Regarding fingolimod (Completeness DQI 8), 26 of the 640 (4.1%)



**Figure 3** Patients with completely missing data at data cut-off (selected Completeness DQIs).

**Notes:** H1 2019, May 1, 2019, data cut-off; H2 2019, November 1, 2019, data cut-off; H1 2020, May 1, 2020, data cut-off; H1 2021, May 1, 2021, data cut-off; H2 2022, November 1, 2022, data cut-off.

**Abbreviations:** DQI, data quality indicator; MS, multiple sclerosis.

discontinued treatment episodes did not have a reason for discontinuation recorded. The participants' date of birth, specifically month and day, was evaluated under Completeness DQI 9. Excepting MSDS3D (which only records year of birth due to local regulations) and one completely missing date of birth in the Swedish MS Registry, all birth dates were complete across MS registries/data sources. Sex was recorded for all patients (Completeness DQI 10). MS onset date (Completeness DQI 11) and MS diagnosis date (Completeness DQI 12) were completely missing for 1.0% and 1.6% of patients, respectively, of the contributing MS registries/data sources (Table 7).

## Discussion

The generation and assessment of CLARION DQIs are critical elements for realizing the full potential and identifying gaps in data that have advanced the understanding of data harmonization and driven the CLARION study toward providing a foundation for the regulatory decision-making process. Importantly, the use of pre-defined DQIs is helping to understand data collection procedures in different MS registries/data sources and identify potential data issues at an early stage, including possible data inconsistencies and methodological differences, thereby allowing for correction and alignment before statistical analyses are performed. Further, the study data can be improved by adapting the extraction, varying the definitions, or excluding sources from specific analysis. The DQI findings to date therefore support the fitness of data for safety evaluation as presented in current and future analyses from this study. Furthermore, the transparent reporting of the generation and conclusions of DQIs will help in the appraisal of the suitability of data for other registry-based studies.

CLARION is a post-approval commitment to both the EMA and the US Food and Drug Administration (FDA). In this regard, a strength of the CLARION study is that the use of DQIs enhances the appropriateness of data usage following the FDA's guidance that real-world data should focus on completeness, consistency, and accuracy, as presented in the FDA's framework on its Real-World Evidence Program (December 2018).<sup>10</sup>

The CLARION study has included patients since the launch of cladribine tablets in the first participating country and now involves numerous MS registries/data sources.<sup>4,5</sup> The target study size has been reached ahead of the inclusion end date, with the enrollment of 8,739 patients by February 2023. Subsequent DQI analysis shows a high level of data consistency and accuracy has been maintained and, in some cases, improved, over time. Improvements in data accuracy and consistency are likely the result of improved recording practices by the clinician, rather than the results of feedback

stemming from previous DQI reports. The experience of the Danish and Swedish MS registries has shown that other data quality metrics, such as the number of yearly visits, are more likely influenced by health national guidelines regarding minimum number of visits a patient should have, which vary greatly between countries participating in CLARION. Of note, the declining proportion of patients with missing reports of lymphocyte count measurements signals an improvement in data quality, an important finding given that this is a key CLARION outcome.<sup>4</sup>

The Representativeness DQIs revealed that the patients' median age at MS onset and diagnosis was generally in the early thirties, and over 90% of included patients had a confirmed diagnosis of relapsing-remitting MS. The female-to-male ratio of 2.6 is within the range estimated in the *Atlas of MS*<sup>11</sup> and consistent with the following: 2.7 in a national prospective population-based observational study conducted in Ireland that identified 292 patients with MS;<sup>12</sup> 2.0 in patients with relapsing-remitting MS identified in a retrospective cross-sectional study conducted in Greece, Switzerland, and Germany,<sup>13</sup> and 2.0 to 2.5 in a previous European registry analysis.<sup>14</sup> The observed distribution of age at MS onset is also similar to previously reported data for patients with MS, and it is associated with several factors, including the completeness and coverage of the data source.<sup>12,14,15</sup> These studies reported a slightly longer time between onset and diagnosis of MS than observed in CLARION, likely reflecting the periods during which these previous studies were conducted (combined 2004 to 2015) and improvements that have allowed for earlier diagnosis in recent years.

There was considerable variability across MS registries/data sources in the proportion of patients who received previous DMTs, possibly reflecting differences in prescribing practices across countries and the completeness of the respective data sources regarding patients with progressive forms of MS and not receiving DMT. Additionally, the application of exclusion criteria that patients with prior fingolimod therapy could not enroll on to cladribine tablet therapy (and vice versa) meant that 16% of patients were excluded (at April 2022 cut-off). Overall, 35.6% and 22.2% of the cladribine and fingolimod cohorts, respectively, were treatment naïve, suggesting that the CLARION population is representative of both treatment-naïve and DMT-experienced patients.

In terms of Consistency DQIs, no clear time-related trends emerged. The number of visits is ultimately influenced by the current health system guidelines and landscape, over which the participating MS registries/data sources have little to no control. For example, the Swedish guidelines state that the aim should be an annual visit to a neurologist, which, in practice, will almost never be to the exact date one year later. Consequently, variability across MS registries/data sources participating in CLARION is to be expected in terms of the number of visits per patient, and consideration must also be given to the coronavirus disease 2019 pandemic that indirectly affected the number of clinical visits. A trend towards personalized management is evident concerning patient visits, with prolonged consultations for those with stable disease and more rigorous management for patients experiencing disease breakthrough.

Accuracy DQIs, assessed by the number of discrepant values, were low, but they improved over time. For discrepant MS treatment stop dates during follow-up, the proportion was low (0–3%) except for the US DoD data source (15.2%). As this is a claims-based database with no stop date recorded as such—proxies were used to define it by adding the number of days of supply to the start date—the higher proportions reported were not unexpected.

Completeness of the data, as summarized in the DQARs over the course of the study to November 1, 2022, remained at similar levels since the first data quality assessment cut-off on May 1, 2019. Basic demographic data (sex, birth date) were complete for all patients in the study excepting one patient in the Swedish MS Registry with a completely missing birth date. Results for DQIs on missing reasons for study treatment discontinuation differ between data sources as the operational definitions used by the data sources for treatment discontinuation differ. The low completeness of end dates for severe lymphopenia suggests that the assessment of repeated severe lymphopenia events is underestimated. For Completeness DQI 2, which assessed records of AESI with missing details and was only evaluated in MSDS3D, 43.9% of AESI records were incomplete. This evaluation, unique to MSDS3D in Germany, was due to the special structure of primary data collection, with multiple modules of electronic case report form data exported (including AESI occurrence, AESI factors, and Medical Dictionary for Regulatory Activities [MeDRA] coding). If, for example, the MeDRA coding was missing for an event, the DQI would be marked as incomplete owing to inconsistency between modules. Regarding Completeness DQIs, only 3.8% patients were lost to follow-up, which highlights the potential of MS registries/data sources to retain patients over time.

The limitations of this study include that not all DQIs could be evaluated in all MS registries/data sources, which affected the interpretation of our findings. In Nordic countries, for example, the evaluation of certain DQIs (eg, records of AESI) requires linkage of MS registries with national registries' data, and this will only be completed for the purposes of interim and final reporting of CLARION. In addition, the definitions for cladribine tablets stop dates vary between MS registries/data sources, largely due to the cladribine treatment regimen, with scheduled stops after each annual treatment course. Evidently, data absence and inconsistencies do not reflect the original registry data; instead, they may have their origin in data manipulation at each registry (as the data management and analysis processes are complex). In CLARION, detailed documentation of registry and study processes and the use of a study-specific common data model for interim and final reports mitigate potential inconsistencies owing to data handling differences between data sources. Further, MS registries/data sources collect data that reflect national clinical practice, adjusted to the developing management and treatment of MS. Collecting data in this fashion may be susceptible to bias, and hence the development and application of the DQIs in CLARION, as has been described.

## Conclusions

The CLARION study shows how the systematic evaluation of 28 DQIs, considering study representativeness, consistency, accuracy, and completeness, allows for the harmonization of data quality and consistency across MS registries/data sources contributing to the post-authorization analysis of safety data. This study presents a unique opportunity to assess how MS registries/data sources complement each other in terms of data quality in absolute terms and relative to each other. In addition, the relevance of the methodology explored in this study could be in its application towards the generation of data from other registry-based studies concerning the quality and consistency demanded by regulatory authorities. Finally, we believe that the analysis suggests that the contributing MS registries/data sources offer data quality for several metrics essential for post-approval safety studies.

## Data Sharing Statement

Any requests for data by qualified scientific and medical researchers for legitimate research purposes will be subject to the data sharing policy of each MS registry/data source.

## Ethics Approval and Informed Consent

The CLARION study is conducted according to Guidelines for Good Pharmacoepidemiology Practice and the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance Code of Conduct, which ensures compliance with General Data Protection Regulation. The study protocol was approved in Germany by the Ethikkommission of the Technische Universität Dresden (reference number: EK 338092018) and in Switzerland by the Ethikkommission Nordwest- und Zentralschweiz (EKNZ) (reference number: 2019-01949). Data permits are granted by approval bodies in each Nordic country.

Informed consent was obtained from all individual participants included in the study.

## Acknowledgments

IQVIA performed all analyses within this report. Medical writing assistance was provided by Caroline Spencer, Claire Snaith, and Steve Winter of inScience Communications, Springer Healthcare Ltd., UK, and supported by Merck.

## Author Contributions

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; agreed to submit to the current journal; gave final approval of the version to be published; and agreed to be accountable for all aspects of the work.

## Funding

The CLARION study is sponsored by Merck (CrossRef Funder ID: 10.13039/100009945).

## Disclosure

**JH** has received honoraria for serving on advisory boards for Biogen, Novartis, and Sanofi, and speaker fees from Bayer, Biogen, BMS, Merck, Novartis, Roche, Sandoz, Sanofi, and Teva. He has served as Principal Investigator for projects, or received unrestricted research support from, Bayer, Biogen, Merck, Sanofi, and Teva. His MS research is funded by the Swedish Research Council and the Swedish Brain foundation. **HB** has received institutional (Monash University) funding from Alexion (Janssen/J&J), Biogen, Merck, and Roche; has carried out contracted research for Alexion (Janssen/J&J), Biogen, Merck, Novartis, and Roche; has taken part in speakers' bureaus for Biogen, Merck, Novartis, and Roche; and has received personal compensation from the Oxford Health Policy Forum for the Brain Health Steering Committee. **MM** has served on scientific advisory boards for AbbVie, Alexion (Janssen/J&J), Biogen, Merck, Novartis, Roche, and Sanofi; has received honoraria for lecturing from Biogen, Merck, Novartis, and Sanofi; and has received research support and support for congress participation from Biogen, Merck, Novartis, Roche, and Sanofi. **SW** has received honoraria for serving on advisory boards for Biogen and Janssen, and speaker fees from Biogen, Janssen, and Novartis. He has served as Principal Investigator for projects from EMD Serono Research & Development Institute, Inc., Billerica, MA, USA, an affiliate of Merck KGaA, and Novartis. **NM** has consulted with Merck and other pharmaceutical companies. **MS-H** has served as an adviser or speaker for Biogen, Celgene (BMS), Novartis, Roche, Sanofi, and Teva; has received institutional research grants for clinical research from Bayer, Biogen, Merck, Novartis, and Roche; and support for congress participation from Biogen, Celgene (BMS), Novartis, Roche, Sanofi, and Teva. **TZ** has served as an advisor or consultant for Biogen, BMS, Merck, Novo Nordisk, Novartis, Roche, Sandoz, Sanofi, Teva, and Viatrix. Served as a speaker or a member of a speaker's bureau for Biogen, BMS, Merck, Novartis, Roche, Sanofi, Sandoz, and Teva. TZ has also received grants for clinical research from Biogen, Novartis, Roche, Sanofi, and Teva. **JK** has received speaker fees, research support, travel support, and/or served on advisory boards by Swiss MS Society, Swiss National Research Foundation (320030\_189140/1), University of Basel, Progressive MS Alliance, Bayer, Biogen, Celgene (BMS), Merck, Novartis, Octave Bioscience, Roche, and Sanofi. **NS** is an employee of Heath Research Tx LLC, whose participation was funded by Merck for this study. **VM** and **IB** are employees of IQVIA, a contract research organization that performs commissioned pharmacoepidemiological studies for several pharmaceutical companies. **MS** is an employee of Merck Healthcare KGaA, Darmstadt, Germany. The authors report no other conflicts of interest in this work.

## References

1. Ziemssen T, Hillert J, Butzkueven H. The importance of collecting structured clinical information on multiple sclerosis. *BMC Med.* 2016;14:81. doi:10.1186/s12916-016-0627-1
2. Glaser A, Butzkueven H, van der Walt A, et al. Big multiple sclerosis data network: an international registry research network. *J Neurol.* 2024;271:3616–3624. doi:10.1007/s00415-024-12303-6
3. Parciak T, Geys L, Helme A, et al. Introducing a core dataset for real-world data in multiple sclerosis registries and cohorts: recommendations from a global task force. *Mult Scler.* 2024;30(3):396–418. doi:10.1177/13524585231216004
4. Butzkueven H, Moore N, Aydemir A, et al. The CLARION study design and status update: a long-term, registry-based study evaluating adverse events of special interest in patients with relapsing multiple sclerosis newly started on cladribine tablets. *Curr Med Res Opin.* 2022;38(7):1167–1176. doi:10.1080/03007995.2022.2059977
5. Butzkueven H, Hillert J, Soilu-Hänninen M, et al. The CLARION study: first report on safety findings in patients newly initiating treatment with cladribine tablets or fingolimod for multiple sclerosis. *Curr Med Res Opin.* 2023;39(10):1367–1374. doi:10.1080/03007995.2023.2256220
6. European Medicines Agency [homepage on the Internet]. Amsterdam: report on multiple sclerosis registries; 2017. Available from: [https://www.ema.europa.eu/en/documents/report/report-multiple-sclerosis-registries\\_en.pdf](https://www.ema.europa.eu/en/documents/report/report-multiple-sclerosis-registries_en.pdf). Accessed March 7, 2023.
7. European Medicines Agency [homepage on the Internet]. Amsterdam: guideline on registry-based studies; 2021. Available from: [https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-registry-based-studies\\_en-0.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-registry-based-studies_en-0.pdf). Accessed June 26, 2023.
8. European Medicines Agency [homepage on the Internet]. Amsterdam: data quality framework for EU medicines regulation; 2023. Available from: [https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/data-quality-framework-eu-medicines-regulation\\_en.pdf](https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/data-quality-framework-eu-medicines-regulation_en.pdf). Accessed January 10, 2024.
9. Ziemssen T, Kern R, Voigt I, Haase R. Data collection in multiple sclerosis: the MSDS approach. *Front Neurol.* 2020;11:445. doi:10.3389/fneur.2020.00445
10. US Food & Drug Administration [homepage on the Internet]. Maryland: framework for FDA's real-world evidence program; 2018. Available from: <https://www.fda.gov/media/120060/download>. Accessed June 26, 2023.
11. The Multiple Sclerosis International Federation (MSIF) [homepage on the Internet]. London: Atlas of MS; 2020. Available from: <https://www.msif.org/wp-content/uploads/2020/10/Atlas-3rd-Edition-Epidemiology-report-EN-updated-30-9-20.pdf>. Accessed April 15, 2024.
12. O'Connell K, Tubridy N, Hutchinson M, McGuigan C. Incidence of multiple sclerosis in the Republic of Ireland: a prospective population-based study. *Mult Scler Relat Disord.* 2017;13:75–80. doi:10.1016/j.msard.2017.02.010

13. Miclea A, Salmen A, Zoehner G, et al. Age-dependent variation of female preponderance across different phenotypes of multiple sclerosis: a retrospective cross-sectional study. *CNS Neurosci Ther.* 2019;25(4):527–531. doi:10.1111/cns.13083
14. Ohle LM, Ellenberger D, Flachenecker P, et al. Chances and challenges of a long-term data repository in multiple sclerosis: 20th birthday of the German MS registry. *Sci Rep.* 2021;11(1):13340. doi:10.1038/s41598-021-92722-x
15. Trojano M, Lucchese G, Graziano G, et al. Geographical variations in sex ratio trends over time in multiple sclerosis. *PLoS One.* 2012;7(10):e48078. doi:10.1371/journal.pone.0048078

Clinical Epidemiology

Dovepress

## Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification, systematic reviews, risk & safety of medical interventions, epidemiology & biostatistical methods, and evaluation of guidelines, translational medicine, health policies & economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Submit your manuscript here: <https://www.dovepress.com/clinical-epidemiology-journal>