Observational long-term safety surveillance registry studies in the Big MS Data (BMSD) network

Study information

Title	Observational long-term safety surveillance registry studies in the Big MS Data (BMSD) network
Protocol version identifier	*V2.0*
Date of last version of protocol	*May 22, 2025*
EU PAS register number	
Active substance	
Medicinal product	
Product reference	
Procedure number	
Marketing authorisation holder(s)	
Joint PASS	*N/A*
Research question and objectives	
Country(-ies) of study	
Author	

Marketing authorisation holder(s)

Marketing authorisation holder(s)	1. Biogen 2. BMS 3. Merck KGaA 4. Novartis 5. Roche 6. Sanofi
MAH Contact Person	

1. Table of contents

2. List of abbreviations

AE	Adverse Event
BMSD	Big MS Data group
CI	Confidence Interval
CIS	Clinically Isolated Syndrome
CNS	Central Nervous System
DMT	Disease Modifying Therapy
EDSS	Expanded Disability Status Scale
EMA	European Medicine Agency
ENCEPP	European Network of Centres for Pharmacoepidemiology and Pharmacovigillance
EUROCAT	European surveillance of congenital anomalies
FDA	Food and Drug Administration
HR	Hazard Ratio
ICD	International Classification of Diseases
IRR	Incidence Rate Ratio
JCV	John Cunningham Virus
LMP	Last Menstrual Period
MAH	Marketing Authorisation Holder
MedDRA	Medical Dictionary for Regulatory Activities
MRI	Magnetic Resonance Imaging
MS	Multiple Sclerosis
PASS	Post Authorisation Safety Study
PY	Person-Years
RWD	Real World Data
SAE	Serious Adverse Event
SOP	Standard Operating Procedure

3. Responsible parties

Marketing Authorisation Holders (MAHs):

1. Biogen
2. BMS
3. Merck KGaA
4. Novartis
5. Roche
6. Sanofi

Due to the dynamics inherent to marketing authorisations and post-approval commitments, MAHs and registries involved in this collaboration may change over time.

Big MS Data (BMSD) Group:

1. Swedish MS Registry (Jan Hillert, Chair of the Swedish MS Registry)
2. Danish MS Registry (Melinda Magyari, Director of The Danish Multiple Sclerosis Registry)
3. Italian Registry (RISM) (Maria Trojano, Chair of the Italian Multiple Sclerosis Registry)
4. French (OFSEP) Registry (Sandra Vukusic, Chair of the French OFSEP Registry)
5. Czech Republic (ReMus) Registry (Jiri Drahota, ReMuS Registry CEO)
6. MSBase International MS Registry (Helmut Butzkueven, Managing Director)

4. Abstract

Title

Observational long-term safety surveillance registry studies in the Big MS Data (BMSD) network

Rationale and background

Approved disease modifying pharmacological treatment for multiple sclerosis (MS) have been used for 30 years and ranged from modestly efficacious and safe to highly efficacious but potentially risky and are typically given for many years making safety assessment in a real world setting attractive.

Research questions and objectives

To assess and characterize the risk of certain safety events in patients with MS, by collecting serious adverse events (SAEs) information in specific disease registries with these objectives:

1) To estimate the incidence/event rates of SAEs

2) To compare the incidence/event rates of specific SAEs between DMTs and comparators

3) Estimate the frequency of adverse pregnancy outcomes in women with MS exposed to DMTs

4) Describe the drug utilization of DMTs

Study Design

Non-interventional descriptive and/or comparative cohort study in MS patients initiating DMT therapy using secondary data derived from MS registries affiliated with the BMSD network. The study cohort will include data from all available MS patients and will estimate rates of SAEs in patients receiving DMT therapies.

Variables

Exposures: Active substance, Route of administration, Date initiated and ceased, Prescribed drug amount or days' supply

Primary outcomes: Serious Adverse Events: Malignancies, Non-melanoma skincancers, Treatment related infections, Other SAEs, All-cause mortality and cause of death.

Secondary outcomes: Pregnancy outcomes and Drug Utilization,

Covariates: Patient characteristics and Disease specific information

Data sources

The Big Multiple Sclerosis Data Network (BMSD) consists of the national MS registries of the Czech Republic, Denmark, France, Italy and Sweden as well as the international MSBase.

Study Size

Of the 300,000 patients in the BMSD approximately 140,000 could be considered for PASS (2024).

Data analysis

The analysis will be conducted using common scripts developed centrally by the study coordinator based on the specific study SAP on data transformed to the common data model (CDM) format using validated appropriate statistical software, primarily R.

Milestones

Reports will be prepared as specified in each post-authorisation safety study (PASS) statistical analysis plan, typically at a yearly basis **.**

5. Amendments and updates

Number	Date	Section of study protocol	Amendment or update	Reason
1	2024-12-05	Full		
2	2025-05-22	Full	<text>	<text>
3	<date>	<text>	<text>	<text>

6. Milestones

Reports will be prepared as specified in each post-authorisation safety study (PASS) statistical analysis plan, typically at a yearly basis. Likewise, interim and final analyses will be defined as part of each product-specific post-authorization safety study (PASS).

7. Rationale

Multiple Sclerosis (MS) is a chronic, inflammatory, and demyelinating disease of the central nervous system (CNS) that affects approximately 2.3 million people worldwide (1). There is currently no cure for MS; however, there are several disease modifying therapies (DMTs) available for its treatment. While the safety and efficacy of DMTs are assessed in clinical trials, these are of relatively short duration and always confined to highly selected patient groups. They specifically exclude women who are pregnant/plan to be pregnant, children, adolescents, older people, and people with significant co-morbidities. Although the number of therapeutic options has increased rapidly in the recent years, real-world data (RWD) on their long-term safety is still scarce. The evaluation of RWD is vital as it offers long-term data collection and is patient rather than product-focused during the lifetime of MS, and thus includes the capacity to capture sequences of treatments throughout patients' disease course.

Registries can capture information from multiple practice settings, including non-trial centers and private practices, and from nations and regions with different DMT exposure characteristics and DMT sequencing due to patient and care team preferences, influence of key opinion leaders in particular jurisdictions and differing DMT approval and reimbursement rules.

The EMA Initiative for Patient Registries recognizes the potential value of using existing MS patient registries to conduct post-authorization studies of safety and effectiveness of MS treatments (2). Although MS patient registries have been extensively used to conduct studies assessing the effectiveness of DMTs, they have until recently not focused on safety. However, all six participating registries recognize the increasing importance of harmonized and consistent safety collection for the benefit of patients. This is particularly important in the modern era where long-acting, highly potent medications will be used and sequenced.

All stakeholders acknowledge that registry-based safety collection will be classified as secondary use of data. Registries will not be subject to any legal requirements for individual case-based adverse event (AE) reporting. This responsibility rests exclusively with the treating physicians according to applicable laws and rules in their jurisdictions.

Information specific to each DMT, including exploration of safety profile or risk management measures that led to the initiation or imposition of a PASS, and a critical review of available published and unpublished data will be outlined separately, in each respective product-specific final PASS protocol with the intention for this to be applied in secondary analyses of data collected according to this general protocol.

8. Research question and objectives

This study aims to assess and characterize the risk of certain safety events in patients with MS (exposed and unexposed to approved DMTs for the treatment of MS), by collecting serious adverse events (SAEs) information in specific disease registries. This may be in response to a public health concern, a risk identified in the risk management plan, an emerging safety issue identified by EMA.

All parts of this core protocol may be subject to customization for each product-specific PASS.

The overall objectives are the following:

1. To estimate the incidence/event rates of SAEs (including non-melanoma skin cancers) in patients with MS exposed to approved DMTs (overall and by individual DMTs)
2. To compare the incidence/event rates of specific SAEs between DMTs and comparators.
3. Estimate the frequency of adverse pregnancy outcomes in women with MS exposed to DMTs during pregnancy or before last menstrual period (LMP)
4. Describe the drug utilization of DMTs.

Further exploratory objectives would be defined in each product-specific PASS.

9. Research methods

9.1. Study design

This PASS is a non-interventional descriptive and/or comparative cohort study in MS patients initiating DMT therapy without prespecified hypotheses. The study will be conducted using secondary data derived from MS registries affiliated with the BMSD network. Disease registries capture observational data from routine clinical practice and the results will therefore reflect prescribing behavior and risk of outcomes in real-world clinical practice.

The study cohort will include all MS patients and will estimate rates of SAEs (including non-melanoma skin cancer) in patients receiving DMT therapies. Additionally, comparative assessments between different DMTs or with unexposed groups and estimation of adverse pregnancy outcomes could also be conducted.

9.2. Setting

9.2.1. Study Population

This study will include ¨patients with MS who are either untreated but eligible for treatment or have initiated treatment with DMTs in routine clinical practice. The study population will be drawn from the registries affiliated with the BMSD Group who fulfill all inclusion and exclusion criteria. All registries will be considered for inclusion; this is expected to cover a broad range of countries. The study population will be defined further in each product-specific protocol.

9.2.2. Selection criteria

This study will include patients who:

1. Have a diagnosis of clinically isolated syndrome (CIS) or MS according to current internationally agreed diagnostic criteria for MS according to the McDonald series of published criteria adopted by the International Advisory Committee on Clinical Trials in Multiple Sclerosis (IACCTMS).
2. Provide informed consent as required by registry and local regulations (unless the data source and project are exempt according to applicable regulatory requirements). Thus, the consent can either be to participate in a specific study or to contribute information to a registry in general, in which case the specific study will be subject to an ethics committee approval.
3. Have at least one administration of the drug of interest

Additional inclusion/exclusion criteria will be defined as part of each product-specific PASS as well as the specific type of MS.

9.2.3. Study Period

The study period (i.e. start of data collection and end of data collection) will be defined in each product-specific protocol. The inclusion period will start after the market launch date in the respective countries of the drug of interest.

The index date will be the day of initiation of the drug of interest and the corresponding date for the comparator group.

9.3. Variables

This section is divided into exposure and outcome variables. Operational definitions for all study variables, including critical variables to meet key study objectives, are provided by the BigMSData network Dictionary and Common Data Model (CDM) which provide variable definitions as well as scripts for transformation of each registry specific data format into the BigMSData CDM.

9.3.1. Exposure

The study drug will be identified in each registry and classified using their Anatomical Therapeutic Chemical (ATC) code.

- The following variables will be extracted from the registries:
    - Active substance
    - Route of administration
    - Date initiated and ceased
    - Prescribed drug amount or days' supply (as appropriate)
- Reason for discontinuation/switch.
- Patients will be classified into two exposure categories: the exposed, and the comparator or standard of care (SOC).

Following the index date, patients will contribute follow-up time and safety event outcomes to their respective exposure category(ies) until they are censored.

Two models for calculating follow-up time will be applied depending on the type of outcome analysed (Figure 1).

Ever exposed model: For each patient follow-up will be calculated, in years, from the index date until the outcome event, death/migration/loss to follow-up or data collection end/study end, whichever occurs first, disregarding any treatment stop or switch to other treatment.

On drug model: For each patient follow-up will be calculated, in years, from the index date until the end of the treatment-specific wash-out period, the outcome event, death/migration/loss to follow-up or data collection end/study end, whichever occurs first. The treatment-specific wash-out period will be calculated for each treatment based on the product characteristics and the literature.

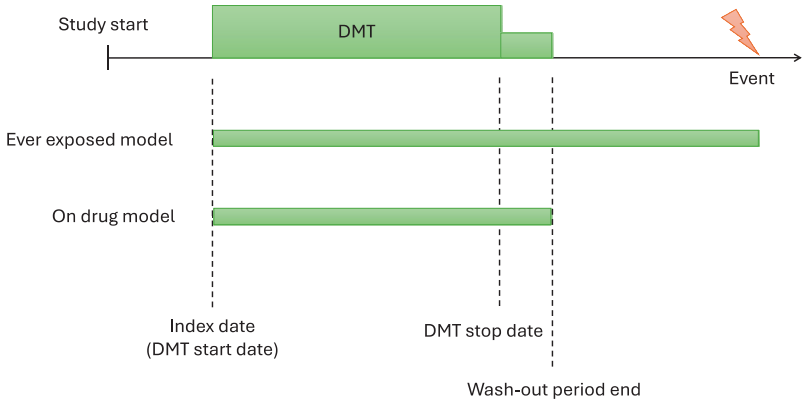

Figure 1. Example of follow-up time calculations for a patient with one disease modifying treatment (DMT) under the "ever exposed" and "on drug" models respectively.

9.3.2. Outcomes

9.3.2.1 Primary outcomes

Serious Adverse Events (including non-melanoma skin cancer)

A serious adverse event (SAE) is any untoward medical occurrence that at any dose results in death, is life-threatening, requires hospitalization or prolongation of existing hospitalization, results in persistent or significant disability or incapacity, or is a congenital anomaly/birth defect [DIR 2001/20/EC Art 2(o)].[3] A SAE does not necessarily have to have a causal relationship with the treatment.

The following SAEs are of Interest:

- Malignancies
- Non-melanoma skin cancers
- Treatment related infections

- Other
- All-cause mortality (if available) and cause of death (if available)

Potential SAEs will be identified using diagnostic codes according to the different coding systems in the registries but will be classified using MedDRA terms whenever possible.

9.3.2.2 Secondary outcomes

Secondary outcomes may be studied in registries which collect those variables or who can access them through linkage with health care data bases.

Pregnancy outcomes

- Date of last menstrual period or other estimated date of pregnancy start
- Ectopic pregnancy (including date of occurrence)
- Spontaneous abortion (pregnancy loss at < 20 weeks) (including date of occurrence)
- Elective termination (including date of occurrence)
    - With foetal defects (classified using EUROCAT) [5].
    - No foetal defects or unknown
- Stillbirth (≥20 completed weeks) (including date of occurrence)
    - With foetal defects (classified using EUROCAT)
    - Without foetal defects or unknown
- Preterm delivery (live birth at < 37 completed weeks) (including date of occurrence)
    - With congenital anomaly (*classified using EUROCAT)*
    - No congenital anomaly or unknown
- Full term birth and outcome (including date of occurrence)
    - With congenital anomaly (*classified using EUROCAT)*
    - No congenital anomaly or unknown

Drug Utilization

Based on the above exposure definitions, the following variables will be derived for the drug utilization patterns:

- Index drug therapy: substance and class of drug therapy, indication duration of use, dosing, route of administration
- Date of discontinuation and reason for discontinuation
- Switching to another therapy
- Number and sequences of successive therapies

9.3.2.3 Covariates

Patient characteristics

- Demographics: age, sex, country of residence
- Socioeconomic information: Employment status (if available)
- Risk factors: Smoking status (never, former, current) and bodyweight (Weight, height) (if available)

- Ethnicity (if available)
- Comorbidities (e.g. none, cardiovascular, respiratory, gastrointestinal, psychiatric, metabolic, malignancies, musculo-skeletal, other auto-immune conditions, other) (if available). Co-morbidities will be be classified using MedDRA terms or other standard coding system (e.g. ICD).
- Medical history (e.g. malignancies) (if available)
- Concomitant medications (if available)

Disease specific information

- Date of MS onset (first clinical manifestation)
- MS type: CIS, Relapsing remitting, Secondary progressive, Primary progressive
- Expanded Disability Status Scale (EDSS) score (including date of assessment)
- Relapses (including date of onset), including glucocorticoid treatment (yes/no)
- JCV antibody status (Pos/Neg with Index) (including date of sample) (if available)
- Date of MS diagnosis (if available)
- MS diagnostic criteria used (if available)
- MRI information (including date of assessment) (if available)
- Lab test results (i.e., lymphocyte counts, liver enzymes), including date of test. (if available)
- Classification can be provided as normal and value if abnormal. (if available)

Important potential confounders to consider – in addition to age, sex, comorbidities, smoking, body weight and a prior history of events of interest – include the duration and severity of the underlying MS. More information on covariates which may be considered in the final analysis due to their potential to confound the association between exposure to medication indicated for MS and the outcomes under the investigation will be detailed in the product-specific PASS protocols.

N.b. Information on covariates is not necessarily captured within registries and therefore marked "(if available)". In some cases, information, e.g. on comorbidities or socioeconomic data, can be obtained by linking to other national or local data sources, which will then be specified by each specific PASS protocol.

9.4. Data sources

European healthcare registries have been previously used to assess the long-term safety in post-marketing commitments requested by the United States Food and Drug Administration (FDA) and EMA. Furthermore, EMA began a patient registry initiative in 2015 to facilitate the use of registry data to inform regulatory decisions (2).

As such, this study focuses on secondary use of data from MS registries affiliated with the BMSD Group network. The Big Multiple Sclerosis Data Network (BMSD) was initiated in 2014 and consists of the national MS registries of the Czech Republic, Denmark, France, Italy and Sweden as well as the international MSBase (3-10).

Table 1. Summary of registries in BMSD

	SMSReg	DMSR	ReMuS	RISM	OFSEP	MSBase
Country	Sweden	Denmark	Czech Republic	Italy	France	International
DMT treatment (ingredient, dosage, start date, end date)	✓	✓	✓	✓	✓	✓
Comorbidities	✓ (through linkage)	✓ (through linkage)	✓	✓	✓ (follow-up) (through linkage)	✓
SAEs	✓ (directly and through linkage)	✓	✓	✓	✓	✓
Death Ascertainment	✓ (through linkage)	✓	✓	✓	✓ (through linkage)	✓
Linkage to other administrative databases	✓ (Swedish National Patient register)	✓ (Danish National Patient register)	*×linkage to be implemented*	*× linkage possible some regions*	(✓) (in process)	*×*

9.5. Study size

All relevant patients with CIS or MS in the BMSD group network will be considered for the purpose of this study. The total number of multiple sclerosis (MS) patients in BMSD amounts to over 300,000 (by end of 2024). Table 2 provides a summary of the number of current patients from each registry and the number of MS patients on active treatment.

Table 2. Number of MS patients by registry

	SMSReg	DMSR	ReMuS	RISM	OFSEP	MSBase
Country	Sweden	Denmark	Czech Republic	Italy	France	International
Total ever MS patients as of extraction date (specify date)	24,224 (2024-11-26)	32,125 (2024-12-02)	22,544 (2024-06-30)	93.802 (2025-04-30)	74,367 (Dec 2022)	84020 (Dec 2024)
Total MS patients on active DMT treatment as of extraction date (specify date)	12,684 (2024-11-26)	10,340 (2024-12-02)	16.355 (2024-06-30)	51.760 (2025-04-30)	37.381 (2022)	48,808 (Dec 2024)

Sample size calculations will be made for study-specific PASS protocols to estimate precision for descriptive studies and power and non-inferiority limits for comparative studies.

9.6. Data management, common data model and data validation

Each contributing registry has its own extensive quality assurance and assessment measures. These include foundational, and intrinsic determinants which are complemented by study/question-specific determinants depending on the specific characteristics of each PASS project. In addition, the Big MS Data Network Common Data Model (CDM) transformation script produces a quality report on each transformed data set which controls for problems in the original registry data as well as certifying that the transformation has been performed correctly.

Performing PASS on MS registry data is an example of secondary data use, as information is collected regardless of an eventual PASS. To quality assure registry data a number of ***foundational determinants of data quality*** are in place for each of the registries, to guarantee quality at the stage of data collection, storage and export. These measures include use of drop-down menus and avoidance of free text format in the IT interfaces, logical blocks and checks to minimize risk of errors, built in checks for data completeness, quality assured data storage, and standardized routines for data export and cleaning.

To improve on both **intrinsic and question specific *determinants of data quality*** used for research purposes within the BMSD network, the participating registries have developed a BMSD common data model (CDM). This creates a common structure that reduces sources of systematic and random errors. The CDM consists of 1) a Data dictionary of agreed upon variables with definitions and descriptions, 2) a script translating each registry's translation of the registry specific data format into the CDM data format 3) a validation script assessing and reporting the data density and completeness of each translated data set. A document describing the BMSD CDM is appended to this protocol.

In addition, routine procedures developed and implemented internally by each participating registry include checking electronic files, maintaining security and data confidentiality, following analysis plans, and performing quality-control checks of all programs. In case of joint studies each registry partner will maintain any patient-identifying information securely on site according to internal/local standard operating procedures or guidance documents. Security processes are in place to ensure the safety of all systems and data. Since these data are of a sensitive nature, every effort is made to ensure that data are kept secure so that they cannot be accessed by anyone except appointed study staff.

Appropriate data storage and archiving procedures are practiced, with periodic backup of files. Standard procedures are in place at each registry to restore files in the event of a hardware or software failure.

9.7. Data analysis

9.7.1. General specifications

A Statistical Analysis Plan (SAP) will be developed for each specific PASS study including the definitions of variables for exposures, outcomes, covariates, and subgroups of interest and then adapted by each registry, which will be based on the BMSD Dictionary and CDM unless deemed inappropriate for the specific design of the drug under study.

The analyses will be conducted using validated appropriate statistical software such as SAS, R, STATA etc. and all code will be annotated using a standardized format.

The analysis will be conducted using common scripts developed centrally by the study coordinator based on the specific study SAP and the CDM. These scripts will implement the same assumptions and methods for variable cleaning and derivation as well as conduct the same statistical analysis to ensure consistency across registries. Fine tuning of the scripts will be done with the help of each registry's teams to ensure adequate model specifications for each data source. The analysis results from each registry will be reviewed by the study coordinator to ensure that the analysis specifications have been met and that the results appear sensible, before combining the results between data sources.

9.7.2. Analysis schedule

Depending on the registry and company-specific protocols, data will be analyzed by registries either individually or in a federated manner, as facilitated by the coordinating center, typically annually. If a protocol specifies federated analysis, details of the data federation model will be outlined in the core SAP for each specific PASS study. Counts of usage of drugs of interest and other variables such as number of SAEs for the specific PASS will be supplied for EMA, as required e.g. annually.

9.7.3. Data handling

To enable sharing of analysis scripts and analysing results, the BMSD Dictionary and CDM tools will be used to convert the registry specific data into the common format as described above. The Dictionary includes the definition of each variable, their format, algorithms to create calculated variables, according to the requirements of the specific PASS study protocol. The code lists used in the registries (ATC, MedDRA, ICD-10) will be implemented within the data dictionary specification form to define the variables and data that will be used in the analysis.

Once study documents have been finalised (and approved by each registry) and data extracted and made available for use, analytical data sets will be created from the anonymised raw patient data. The data handling will include the creation of variables defined in the protocol, quality checks and data cleaning:

- Choice of variables: the data dictionary defines the levels of each variable (e.g., gender: 1=male; 2=female); it also defines the time frames of reference for each variable. All programming will be documented.

- The CDM tool documents the success of transformation and creates a report on data quality, density and eventual conflicts. Remaining queries or issues with the converted data will be documented and resolved.

The resulting analytical datasets will be used for the analysis in a meta-analytical fashion.

9.7.4. General presentation of summaries and analyses

Continuous variables will be summarised with the statistics mean, standard deviation (SD), median, 25% quartile (Q1), 75% quartile (Q3), minimum and maximum. Mean, SD, median, Q1 and Q3, minimum and maximum values, where available, will be displayed to the same level of precision as collected in the respective data source, and all other continuous variables created will be to 1 decimal place, unless stated otherwise. Categorical variables will be summarised as frequencies and percentages. Additionally, as per general reporting guidelines, no cells containing <5 events, or results based on <5 events, will be reported to avoid the possibility of unintentional (deductive) disclosure. A missing category will be included as applicable, or the sample size (patients with non-missing results) will be displayed. Presentation of percentages will be to 1 decimal place.

Unless stated otherwise, all confidence intervals will be 2-sided and conducted at the 95% confidence level.

For descriptive analyses, missing data will be omitted from the percentages, but the frequency will be presented as a separate category for categorical data and the amount of missing data for continuous variables will be summarized. Imputation of missing data may be considered if indicated.

Incidence rates per 100 or 1,000 person years with two-sided 95% confidence intervals (CI) will assume a Poisson-type distribution including, but not limited to, negative binomial, zero-

inflated Poisson and/or conventional Poisson as appropriate given the underlying distribution of the outcome or event being summarised. Presentation of incidence rates, incidence rate ratios and other effect measures will be to 2 decimal places. Unless stated otherwise, all statistical tests will be 2-sided and conducted at the 0.05 alpha level. P-values will be presented to 3 decimal places.

Multiple events will be expressed as cumulative incidence rates per 100 or 1,000 person years with two-sided 95% confidence intervals (CI) assuming a Poisson-type distribution as described above.

Associations of binary outcome variables will be described using odds ratios and 95% confidence intervals. Time-to-event variables will be described with Kaplan-Meier curves and median time to event will be described. Associations will be calculated by survival techniques such as Cox proportional hazards regression models or marginal Cox models as indicated. Linear regression will be used to describe associations of continuous normal outcomes. Non-linear regression variants will be considered in cases of significant non-linearity in continuous endpoints. Point estimates and 95% confidence intervals will be presented. Model assumptions and goodness-of-fit will be tested for all regression-based analysis.

Risk/rate differences and number needed to harm, and 95% confidence intervals, will be calculated, where appropriate

Drug utilization patterns will be displayed using visualization techniques, such as Sankey charts or sunburst diagrams, etc, as appropriate.

9.7.5 Study-specific analyses

Study-specific analyses will be tailored to the objectives of the specific PASS objectives which may be descriptive to evaluate rates of occurrence of SAEs or drug utilization, and/or comparative to evaluate rates of occurrence of SAE between different drugs or groups of drugs or standard of care.

Unadjusted and adjusted (by age, sex and other covariates) rates and rate ratios (incidence, odds, hazards, relative risks, risk difference) will be calculated, as appropriate.

Comparative analyses may use various methods to ensure comparability and validity including, but not limited to, propensity score adjustment. The target trial approach may be adopted in the design of the analysis.

To adjust for confounding, various methods such as covariable adjustment and various propensity score methods (matching, stratification, inverse-probability weight) may be used, as appropriate to the specific PASS.

Immortal time bias between treatment groups will be accounted for by analysis of treatment as a time-updated covariate or addressed in the analytical design, such as updated time-varying covariate adjustment and/or pairwise censoring, as appropriate to the specific PASS.

9.7.6. Subgroup analyses

For each subgroup defined in a categorical variable will be created indicating which stratum a patient belongs to. Summary statistics will be calculated as described as above stratified by the subgroups. For each outcome the same analytical model will be used but also include the subgroup variable in the model. Stratum specific effect sizes (e.g., incidence rates, odds ratios, hazard ratios or differences in means) will be calculated along with 95% confidence intervals. Interaction tests will be employed on subgroup analyses.

9.7.7. Missing data

Missing data is anticipated in relatively few key variables (see Table 3). Missing data will be recorded as such (e.g., coded as 99), but with those records retained or excluded as appropriate. Missing data may affect dates, baseline assessments, and other follow-up variables. For descriptive studies no imputation is foreseen, and the proportion of missing values will be reported.

For comparative analyses, where the proportion of missingness is sufficiently low (e.g. <25%), a complete case analysis will be conducted. After documenting the nature, pattern, and quantity of missing data, further techniques to handle missing data may be considered such as imputation techniques including, but not limited to, multiple imputation via chained equations.

9.7.8. Sensitivity analyses

Sensitivity analyses are PASS-study specific and will be developed for each PASS. These may affect the misclassification of exposure (e.g. estimates from a validation sub-study), outcome (e.g. estimate of positive predictive value form a validation sub-study) and by levels of adherence. Potential time-related biases may be explored by analyses by time and adjusting for key time-varying covariates.

The E-value will be calculated and is defined as the minimum strength of association, on the risk ratio scale, that an unmeasured confounder would need to have with both the treatment and the outcome to fully explain away a specific treatment-outcome association, conditional on the measured covariates.

9.7.9. Pooled analyses

If the study populations included in each registry meet requirements for homogeneity, techniques will be used to pool the aggregate data from the different registries to produce overall effect estimates. The pooled analysis will be designed to estimate the effect of the exposure while controlling for confounding. Each registry will be retained as a stratification variable or independent analysis, so the individual estimates within each registry can be assessed. Each registry may be modelled as a random effect for the analysis of pooled data to adjust for unobserved or unquantified heterogeneity between and within each registry. There will be an exploration of reasons for any statistical heterogeneity, for example by comparing patient populations and incident rates between registries. Registry-specific incidence rates, rate ratios and drug utilization rates will be presented as Forest plots.

Mantel-Haenszel methods for meta-analysis will be used to pool the aggregate data from each registry and calculate overall adjusted incidence rate ratios.

Meta-analysis will be conducted using random effects as the default primary analysis. A fixed effects meta-analysis may be conducted as a sensitivity analysis if indicated. Heterogeneity between registries will be measured using I^2 index, which represents the percentage of variability in results across registries that is due to real differences rather than due to chance. An I^2 of less than 25% will be viewed as low heterogeneity, between 25% and 50% as moderate, and over 50% as high heterogeneity. Cochran's Q test will also be performed as a formal statistical test of heterogeneity.

Meta-analyses will be conducted for the following analyses:

- Pooled incidence rates for SAEs in the drug of interest and, where relevant, comparator cohorts
- Pooled all-cause mortality rate in the drug of interest and, where relevant, comparator cohorts
- Pooled rates of utilization of the drug of interest.

9.8. Quality control

Each registry will use its own standard operating procedures, internal policies and process guidance. This study will be conducted as part of routine clinical practice. The Standard Operating Procedures (SOPs) may include rules for data storage, methods to maintain and archive project and study documents, quality-control procedures for programming, standards for writing analysis plans, review of analysis programs and study documents by senior staff as well as internal audits if applicable. Registries will seek relevant certification either individually or as a group (e.g. EMA certification).

Independent review of results may be made available to external reviewers in connection with a specific PASS.

For the sake of assuring quality of data and results of analyses measures have been taken to assure the five dimensions of data quality specified in the ***Data Quality Framework for EU medicines regulation*** publication (30 October 2023, Data Analytics and Methods Task Force, EMA/326985/2023). In summary, data ***reliability***, including data ***accuracy*** and ***precision***, is assured by the foundational determinants described above and specified in the Request Tool document and appendices. The data ***coherence*** is assured by applying the BMSD CDM which promotes format, structural and semantic consistency of contributing data sets. ***Timeliness*** is guaranteed by the respective registry specific operational procedures described in the Request Tool. Data ***relevance*** for the purpose of performing PASS has been indicated by the results of formal ***feasibility study*** report showing each registry's ability to collect safety data.

Finally, data ***extensiveness*** is promoted by i) the size of the patient population of the BMSD registries, together representing around 10 % of the prevalent world MS population, ii) by the high population coverage of the contributing national registries and iii) by the completeness of data among patients relevant for inclusion in a future PASS.

Thus, to give an example of data extensiveness of relevant registry data, BMSD has defined cohorts of patients from each contributing registry who are active in the sense that they are on treatment and have been seen by a documented visit in the past 36 months. Table 3 shows data density of the key clinical variables required for PASS by the end of 2024.

Table 3. Characteristics and data density of registry populations relevant for the performance of PASS, defined as patients on treatment with data recorded from a visit within the past 36 months (data by the end of 2024):

Seen & treated past 36 months

	SMSreg	DMSR	REMUS	OFSEP	RISM	MSBase
N	11181	10903	17093	25597	34232	40790
Female %	69.10%	68.8 %	70.5 %	72.2 %	66.9 %	71.0 %
Mean age (SD)	45.64 (11.79)	47.9 (11.8)	45.5 (11.2)	46.6 (12.1)	47.3 (12.35)	46.0 (12.2)
Documented age	100.00 %	100.0 %	100.0 %	100.0 %	100 %	100.0 %
Documented sex	100.00 %	100.0 %	100.0 %	100.0 %	100 %	100.0 %
Documented MS onset date	98.07 %	100.0 %	100.0 %	100.0 %	98.3 %	98. 5%
Documented diagnosis date	97.76 %	99.2 %	84.1 %	83.2 %	96.3 %	77. 0%
Documented MS course	99.69 %	100.0 %	98.8 %	100 %	95.3 %	96.1 %
Ever recorded relapses	60.41 %	77.8 %	98.6 %	100 %	79.2 %	80.8 %
>= one MRI	93.39 %	99.5 %	75.1 %	97.5 %	93.1 %	87.2 %
>= two MRI	88.95 %	97.3 %	68.0 %	95.7 %	90.9 %	75.7 %
>= one EDSS	94.99 %	99.9 %	100.0 %	100.0 %	100 %	98.9 %

9.9. Limitations of the research methods

This study will have the main limitations typical of secondary use of data studies (e.g. sample size dependent of medication uptake in clinical practice, missing data, channeling bias, residual confounding, etc.). Further details of potential limitations will be outlined in each product-specific PASS final protocol.

9.10. Other aspects

None.

10. Protection of human subjects

This is a non-interventional study using secondary data. All data used in the study will be de-identified with no breach of confidentiality with regard to personal identifiers or health information.

The study will be conducted in agreement with the regulation (EU) 2016/679 of the European Parliament on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (GDPR). Data protection and privacy regulations will be observed in collecting, forwarding, processing, and storing data from study participants.

Regulatory and ethical requirements will be followed in each country where the respective registries are used. Each registry will apply for an independent ethics committee review, if required according to local regulations. The study will comply with the module VIII of the good pharmacovigilance practices (GVP).

All registries have obtained informed consent for data to be used for PASS (and other non-interventional studies) that collect only observational data.

11. Management and reporting of adverse events/adverse reactions

For non-interventional studies that are solely based on secondary use of data, reporting of adverse events/adverse drug reactions is not required. As per the EMA Guideline on GVP (Module VI–Management and reporting of adverse reactions to medicinal products [Revision 1]), individual reporting of adverse reactions is not required for non-interventional study designs that are based on secondary use of data (11). The responsibility for individual case-based adverse event (AE) reporting rests exclusively with the treating physicians according to applicable laws and rules in their jurisdictions and registries are therefore not subject to any legal requirements.

Reports of adverse events/adverse drug reactions should only be summarized in the study report, where applicable (12, 13).

12. Plans for disseminating and communicating study results

Regardless of the outcome of study, all MAHs are dedicated to openly providing information on the results to healthcare professionals and to the public. This study will be registered in the EU PAS register and the EMA approved protocol posted on this website before analysis begins. A study report will be posted after completion of the study (14).

Progress reports will be prepared over the data collection period as agreed with the relevant MAHs and submitted by each respective MAH to EMA through scheduled regulatory safety reporting. The interim and final study reports will accurately and completely present the study objectives, methods, results, limitations, and interpretation of the findings and be sent to the EMA. The final report will follow the structure recommended in the EMA template "Guidance for the format and content of the final study report of non-interventional post-authorisation safety studies" (15).

Following approval of the study report by EMA, the results will be presented in appropriate scientific conferences and in peer-reviewed journals.

13. References

1. Multiple Sclerosis International Federation (MSIF). Atlas of MS 2013: Mapping Multiple Sclerosis Around the World. 2013.
2. European Medicines Agency (EMA). "Report in Multiple Sclerosis Registries - Workshop 7 July 2017," 2017.
3. Butzkueven H, Chapman J, Cristiano E, Grand'Maison F, Hoffmann M, Izquierdo G, et al. MSBase: an international, online registry and platform for collaborative outcomes research in multiple sclerosis. Mult Scler. 2006;12(6):769-74.
4. Flachenecker P, Buckow K, Pugliatti M, Kes VB, Battaglia MA, Boyko A, et al. Multiple sclerosis registries in Europe - results of a systematic survey. Mult Scler. 2014;20(11):1523-32.
5. Glaser A, Stahmann A, Meissner T, Flachenecker P, Horáková D, Zaratin P, et al. Multiple sclerosis registries in Europe - An updated mapping survey. Mult Scler Relat Disord. 2019;27:171-8.
6. Hillert J, Stawiarz L. The Swedish MS registry – clinical support tool and scientific resource. Acta Neurol Scand. 2015;132(199):11-9.
7. Horáková D. Jaká data nabízí celostátní registr pacientů s roztroušenou sklerózou ReMuS? 2020.
8. Magyari M, Joensen H, Laursen B, Koch-Henriksen N. The Danish Multiple Sclerosis Registry. Brain Behav. 2021;11(1):e01921.
9. Trojano M, Bergamaschi R, Amato MP, Comi G, Ghezzi A, Lepore V, et al. The Italian multiple sclerosis register. Neurol Sci. 2019;40(1):155-65.
10. Vukusic S, Casey R, Rollot F, Brochet B, Pelletier J, Laplaud DA, et al. Observatoire Français de la Sclérose en Plaques (OFSEP): A unique multimodal nationwide MS registry in France. Mult Scler. 2020;26(1):118-22.
11. EMA GVP. Module VI - Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2); EMA/873138/2011_Rev 2, 28 Jul 2017.
12. European Medicines Agency and Heads of Medicines Agencies, "Guideline on good pharmacovigilance practices (GVP) - Annex I - Definitions (Rev 4)," 2017. .
13. European Medicines Agency and Heads of Medicines Agencies, "Guideline on good pharmacovigilance practices (GVP) - Module VI – Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2)," 2017.
14. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). The European Union electronic Register of Post-Authorisation Studies (EU PAS Register) [Available from: https://www.encepp.eu/encepp_studies/indexRegister.shtml.
15. European Medicines Agency (EMA). Guidance for the format and content of the final study report of non-interventional post-authorisation safety studies. 2013.

Annex 1: List of stand-alone documents

Number	Document reference number	Date	Title

Annex 2: ENCePP Checklist for Study Protocols

Study title: Prospective observational long-term safety surveillance registry studies in the Big MS Data (BMSD) Group network

EU PAS Register® number: **Study reference number (if applicable):**

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	9.2
1.1.2 End of data collection[2]	☒	☐	☐	9.2
1.1.3 Progress report(s)	☒	☐	☐	12
1.1.4 Interim report(s)	☒	☐	☐	12
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	12
1.1.6 Final report of study results.	☒	☐	☐	12

Comments:

As this is a generic protocol, the details will be described and elaborated in each specific PASS protocol

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1 Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	8
2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	8
2.1.2 The objective(s) of the study?	☒	☐	☐	8
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	8, 9.2
2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☐	☒	
2.1.5 If applicable, that there is no *a priori* hypothesis?	☒	☐	☐	8

Comments:

As this is a generic protocol, the details will be described and elaborated in each specific PASS protocol

[1] Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.
[2] Date from which the analytical dataset is completely available.

Section 3: Study design	**Yes**	**No**	**N/A**	**Section Number**
3.1 Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	☒	☐	☐	9.1
3.2 Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	9.1
3.3 Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	9.7.4
3.4 Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☒	☐	☐	9.7.5
3.5 Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☒	☐	☐	11

Comments:

As this is a generic protocol, the details will be described and elaborated in each specific PASS protocol

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.1 Is the source population described?	☒	☐	☐	9.2
4.2 Is the planned study population defined in terms of:				
4.2.1 Study time period	☒	☐	☐	9.2
4.2.2 Age and sex	☒	☐	☐	
4.2.3 Country of origin	☒	☐	☐	
4.2.4 Disease/indication	☒	☐	☐	
4.2.5 Duration of follow-up	☒	☐	☐	
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	9.2

Comments:

As this is a generic protocol, the details will be described and elaborated in each specific PASS protocol

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.1 Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	9.3.1
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☐	☐	☒	

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.3 Is exposure categorised according to time windows?	☐	☐	☒	
5.4 Is intensity of exposure addressed? (e.g. dose, duration)	☐	☐	☒	
5.5 Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☐	☒	
5.6 Is (are) (an) appropriate comparator(s) identified?	☒	☐	☐	

Comments:

As this is a generic protocol, the details of the exposure will be described and elaborated in each specific drug-specific PASS protocol

Section 6: Outcome definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
6.1 Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☒	☐	☐	9.3.2
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	9.3.2
6.3 Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☒	☐	☐	9.3.2
6.4 Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☐	☐	☒	

Comments:

This is a core protocol specific for PASS studies however this protocol does describe disease management and drug utilization.

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☒	☐	☐	9.3.2
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☒	☐	☐	9.7.6
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	9.7.6, 9.7.7, 9.7.8

Comments:

Section 8: Effect measure modification	**Yes**	**No**	**N/A**	**Section Number**
8.1 Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☒	☐	☐	9.7.6

Comments:

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	9.4
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	9.4
9.1.3 Covariates and other characteristics?	☒	☐	☐	9.4
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	9.3.1
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	9.3.1
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	9.3.1
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	9.3.1
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	9.3.2
9.3.3 Covariates and other characteristics?	☒	☐	☐	9.3.2.3
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☐	☐	☒	

Comments:

As this is a generic protocol, the details of the linkage methods will be described and elaborated in each specific drug-specific PASS protocol.

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	9.7
10.2 Is study size and/or statistical precision estimated?	☒	☐	☐	9.5

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.3 Are descriptive analyses included?	☒	☐	☐	9.7
10.4 Are stratified analyses included?	☒	☐	☐	9.7.6
10.5 Does the plan describe methods for analytic control of confounding?	☒	☐	☐	9.7.5
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☒	☐	☐	9.7.8
10.7 Does the plan describe methods for handling missing data?	☒	☐	☐	9.7.7
10.8 Are relevant sensitivity analyses described?	☒	☐	☐	9.7.8

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	9.8
11.2 Are methods of quality assurance described?	☒	☐	☐	9.8
11.3 Is there a system in place for independent review of study results?	☒	☐	☐	9.8

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☒	☐	☐	9.9
12.1.2 Information bias?	☒	☐	☐	9.9
12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☒	☐	☐	9.9
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☒	☐	☐	9.5

Comments:

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	10
13.2 Has any outcome of an ethical review procedure been addressed?	☐	☐	☒	
13.3 Have data protection requirements been described?	☒	☐	☐	10

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	5

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	12
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	12

Comments:

Annex 3: Known limitations of primary variable collection and procedures

Variable not collected/incomplete:	**In registry:**
JCV status and index	OFSEP registry
Date and course of death	Will be incomplete because the MS centre may not be notified. Available as complete data through data linkage from Sweden and Denmark.
Big MS data merge	The Italian registry may agree to merge safety data annually or biannually under existing BMSD procedures and governance rules, but only aggregated results of this analysis may be provided to pharma companies.