NOVARTIS

Chief Medical Office and Patient Safety

Non-Interventional Study Protocol (PASS) with secondary use of data

REDACTED PROTOCOL

COMB157G2406

Title	Kesimpta long-term retrospective safety study utilizing real-world data from existing multiple sclerosis registries and databases from multiple countries
Protocol version identifier	V1.0
Date of last version of protocol	17 March 2023
EU PAS register number	Study not registered
Active substance	Ofatumumab (ATC code: L04AA52)
Medicinal product	Kesimpta
Product reference	EMEA/H/C/005410
Procedure number	Not applicable
Name of Marketing authorization holder(s)	Novartis Ireland Ltd Vista Building Elm Park Merrion Road, Ballsbridge, Dublin 4, D04 A9N6
Joint PASS	No

Research question and objectives	The research question: • Is Kesimpta initiation at any time, dose and/or duration during a multiple sclerosis (MS) patient's treatment journey associated with an increased risk of malignancy and infections compared to other disease modifying therapies (DMTs) initiation? Primary objective: • In patients diagnosed with MS, to compare the risk of 1) malignancy (except non melanoma skin cancers [NMSC] overall and for pre-defined type) and 2) late-onset infections (overall, by type and seriousness) between Kesimpta-initiators and other DMT-initiators irrespective of therapy discontinuation or switch and 3) acute-onset infections (overall, by type and seriousness) between Kesimpta-initiators and other DMT-initiators while on therapy. Secondary objectives: • To characterize Kesimpta and other DMT-initiator populations and Kesimpta and other DMT use including but not limited to patient demographics, MS disease characteristics, duration of the Kesimpta and other DMT use, number of treatment switches • To estimate the incidence rates of 1) malignancies (overall, except NMSC and for pre-defined type) and 2) late-onset infections (overall and by type) following Kesimpta and other DMT initiation irrespective of therapy discontinuation or switch and 3) acute-onset infections (overall and by type) following Kesimpta and other DMT initiation while exposed to Kesimpta or other DMT • To assess long term safety, estimate and compare the risk of serious adverse events between Kesimpta-initiators and other DMT-initiators while on therapy
Country (-ies) of study	Planned: Denmark, Sweden, France, Italy and MSBase (multi country data source)

Author [redacted] (MSc, [redacted])
[redacted] (MD, MPH, PhD, [redacted])
[redacted] (PhD, [redacted])
[redacted] (MSc, [redacted])
[redacted] (PhD, [redacted])
[redacted] (MSc, [redacted] Novartis AG)
[redacted] (PhD, [redacted], [redacted], Novartis Pharmaceuticals Corporation)

Marketing authorization holder(s)

Marketing authorization holder(s)	Novartis Ireland Ltd Vista Building Elm Park Merrion Road, Ballsbridge, Dublin 4, D04 A9N6
MAH contact person	[redacted] Regulatory Affairs Novartis Pharma AG Fabrikstrasse 27, Novartis Campus CH-4056 Basel, Switzerland

Table of contents

List of figures

List of tables

List of abbreviations

AE	Adverse Event
ATC	Anatomical Therapeutic Chemical
DALY	Disability-Adjusted Life Year
DMT	Disease Modifying Therapy
EC	Ethic Committee
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU	European Union
FDA	Food & Drug Administration
GPP	Good Pharmacoepidemiology Practices
GVP	Good Pharmacovigilance Practices
HA	Health Authority
HBV	Hepatitis B virus
HR	Hazard ratio
ICMJE	International Committee of Medical Journal Editors
ID	Index date
ISPE	International Society for Pharmacoepidemiology
MAH	Marketing Authorization Holder
MedDRA	Medical Dictionary for Regulatory Activities
MS	Multiple Sclerosis
NI	Non-Interventional
NIS	Non-Interventional Study
NMSC	Non-melanoma skin cancer
PAS	Post-Authorization Study
PASS	Post-Authorization Safety Study
PI	Principal Investigator
PIC	Principle-in-Charge
PML	Progressive Multifocal Leukoencephalopathy
PPMS	Primary-Progressive Multiple Sclerosis
PRMS	Progressive-Relapsing Multiple Sclerosis
PY	Person-years
QALY	Quality-Adjusted Life-Year
RRMS	Relapsing-Remitting Multiple Sclerosis
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SID	Study Identification Number
SPMS	Secondary Progressive Multiple Sclerosis
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
WHO	World Health Organisation

1 Responsible parties

Table 1-1 Responsible parties

Role	Person
MAH study lead	[redacted] Novartis Pharmaceuticals Corporation One Health Plaza, East Hanover, NJ 07936-1080, US Email: [redacted]
Coordinating center	[redacted]
Principal-in-Charge (PIC)	[redacted], B.Sc. Pharmacy
Principal investigator (PI)	[redacted], MD, MPH, PhD
Main protocol author	[redacted], MSc
Other protocol authors	[redacted], MD, MPH, PhD [redacted], PhD [redacted], MSc [redacted], PhD [redacted], Novartis Pharmaceuticals Corporation One Health Plaza, East Hanover, NJ 07936-1080, US

	[redacted] Novartis Pharma AG Fabrikstrasse 27, Novartis Campus CH-4056 Basel, Switzerland
Investigators who have substantially contributed to the study conception or design	[redacted]
MAH, marketing authorization holder	Novartis Ireland Ltd Vista Building Elm Park Merrion Road, Ballsbridge, Dublin 4, D04

2 Abstract

Title

Kesimpta long-term retrospective safety study utilizing real-world data from existing multiple sclerosis registries and databases from multiple countries

Version and date

V1.0, 17 March 2023

Name and affiliation of main author

[redacted]

MSc, [redacted]

Rationale and background

Kesimpta (ofatumumab), a fully human monoclonal antibody that binds to human cluster of differentiation 20 (CD20) molecules expressed on B-cells, is approved in the United States (US) and in Europe for treating adults with relapsing forms of multiple sclerosis (RMS). Currently there are limited data on the risk of infections and long-term effects on the risk of malignancies following exposure to Kesimpta. The study is designed to assess whether exposure to Kesimpta increases the risk of malignancies and infections.

Research question and objectives

Is Kesimpta initiation at any time, dose and/or duration during a multiple sclerosis (MS) patient's treatment journey associated with an increased risk of malignancy and infections compared to other disease modifying therapies (DMTs) initiation?

Primary objective:

- In patients diagnosed with MS, to compare the risk of 1) malignancy (except non melanoma skin cancers [NMSC] overall and for pre-defined type) and 2) late-onset infections* (overall, by type and seriousness) between Kesimpta-initiators and other DMT-initiators irrespective of therapy discontinuation or switch and 3) acute-onset infections* (overall, by type and seriousness) between Kesimpta-initiators and other DMT-initiators while on therapy.
  *including opportunistic infections

Overall secondary objectives:

- To characterize Kesimpta and other DMT-initiator populations and Kesimpta and other DMT use including but not limited to patient demographics, MS disease characteristics, duration of the Kesimpta and other DMT use, number of treatment switches.
- To estimate the incidence rates of 1) malignancies (overall except NMSC and for pre-defined type) and 2) late-onset infections (overall and by type) following Kesimpta and other DMT initiation irrespective of therapy discontinuation or switch and 3) acute-onset infections (overall and by type) following Kesimpta and other DMT initiation while exposed to Kesimpta or other DMT.
- To assess long term safety, estimate and compare the risk of serious adverse events (SAEs) between Kesimpta-initiators and other DMT-initiators while on therapy.

Study design

The study is an observational, comparative, retrospective, new user cohort study using longitudinal secondary data from national and multi-national real-world databases.

Patients with MS are assessed for exposure to Kesimpta and other DMT and followed-up for malignancies and infections. Outcomes are compared between Kesimpta- and other DMT-initiators.

The target is to follow, for a minimum of 5 years, a minimum of 3000 Kesimpta-initiators and to compare these to other DMT-initiators during the same period. The study duration is planned for 8 years post Kesimpta launch in the database countries including a 3 year-accrual period.

This is a Master protocol. Localized study protocols will provide further study details for each database/country.

Setting and study population

The study population consists of adult patients with MS initiating Kesimpta or another DMT after Kesimpta's launch date in the target countries. Patients younger than 18 years or exposed to Kesimpta or other DMT for non-MS indications are excluded. For the other DMT-initiator cohort, patients with prior Kesimpta use are excluded.

Variables

The following variables are included:

- Exposure variables: Kesimpta or other DMT initiations during study period
- Outcomes:
    - Primary outcomes: malignancies (excluding NMSC), pre-defined malignancies, late and acute-onset infections (overall, by type and seriousness) including opportunistic infections.
    - Secondary outcomes: serious adverse events (SAE) (overall, by type, if feasible), suicidal ideation, intestinal or bowel obstruction and sarcoidosis (if feasible)
- Note: outcomes will be classified in two categories requiring distinct analytical conventions. Malignancies, late-onset infections and sarcoidosis will be handled as "Long-term outcomes"; while SAEs, acute-onset infections, suicidal ideation and intestinal or bowel obstruction will be handled as "Short-term outcomes".
- Censoring variables: Death, end of data availability (emigration, end of follow-up in the data, or analysis cut-off date)
- Other variables: Characteristics related to demographics, lifestyle, MS disease, Medical history, comorbidities and comedications.

Data sources

This study plans to utilize data from data sources in Denmark, Sweden, France, Italy and MSBase, selected based on the outcome of a feasibility assessment conducted in Q4 2021.

Study size

The follow-up is planned to last for a minimum of 5 years. At least 3000 Kesimpta-initiators are planned to be accrued over a 3-year period across all databases.

The meta-analysis design and switching of patients between treatment cohorts are ignored in the following sample size calculations. Due to the observational nature of this study, neither the size of the Kesimpta cohort nor the ratio vs. comparator can be known, therefore several scenarios are provided to illustrate the range of hazard ratios (HR) that can be ruled out under a given set of parameters.

Assuming a 4.08 per 1000 Person-years (PY) incidence rate for malignancy (overall except NMSC) (Norgaard et al, 2021), a drop-out rate of 5% per year and no difference in hazard rates between Kesimpta and other DMT-initiators, the expected effect size (expressed in terms of HR) that can be ruled out with a N=3000 Kesimpta cohort size and Kesimpta:Comparator (K:C) ratio of 1:1, 1:3 and 1:10 are 1.61, 1.48 and 1.43 respectively. For colorectal and cervical cancers and any other study outcomes with incidence rates below 1.00/1000 PY, it will be difficult to show that there is no difference between the cohorts. For infection related outcomes with incidence above 20.00/1000 PY, HR less than 1.25 can be ruled out.

Data analysis

Analyses will be performed in two stages. First, data will be analyzed locally for each data source following a common statistical methodology. Second, the aggregated data or stratified summaries, as appropriate, from each data source will be provided to [redacted] to conduct integrated analyses using meta-analytical methods.

Annual update reports will be descriptive. Final and interim reports will include both descriptive and comparative analyses.

In each data source, the Kesimpta-initiator and other DMT-initiator cohorts will be extracted and described in terms of patient demographics, potential confounders for malignancy and infections, drug use and duration of follow-up. In addition, for each study outcome of interest, the total number of incident and recurrent events, cumulative person time and unadjusted incidence and event rates with 95% confidence intervals (CIs) will be presented.

- **For malignancies, late-onset infections and sarcoidosis**, a "treatment policy" strategy will be adopted, and follow-up will end at the first event occurrence, or when the patient is censored from the database (e.g., due to emigration, analysis cut-off date) whichever occurs first.
- **For acute-onset infections, SAEs, suicidal ideation and intestinal or bowel obstruction**, a "while on treatment" strategy will be adopted and follow-up will end at treatment discontinuation + drug specific wash-out period days, or when the patient is censored from the database (e.g., due to emigration, analysis cut-off date) whichever occurs first.

In interim and final reports, time to event of each outcome of interest will be compared between the Kesimpta and other DMT-initiator cohorts using Cox proportional-hazard model. Unadjusted and adjusted HR along with 95% CIs will be reported.

To account for the non-randomized nature of the comparison, propensity score (PS) methods (accounting for but not limited to a pre-specified list of confounders) will be used to ensure cohort comparability at index date (Austin 2011). Overlap in the PS distribution across cohorts will be assessed.

Milestones

Planned dates of study milestones:

Start of data collection, i.e. start date of data extraction: Jan 2025

End of data collection, i.e. date from which the analytical dataset is completely available: Jul 2032

Annual Update Report 1: 30 Oct 2025

Annual Update Report 2: 30 Oct 2026

Annual Update Report 3: 30 Oct 2027

Annual Update Report 4: 30 Oct 2028

Annual Update Report 5/Interim report 1: 30 Oct 2029

Annual Update Report 6: 30 Oct 2030

Annual Update Report 7: 30 Oct 2031

Registration in the EU PAS register: within 30 days after protocol endorsement by EMA

Final Extract: Jul 2032

Final report of study results: 30 Oct 2033

3 Amendments and updates

None.

4 Milestones

The annual interim reports and the final report will be submitted as standalone procedures.

Table 4-1 Planned dates of study milestones

Milestone	Planned date[1]
Start of data collection, i.e., start date of data extraction	Jan 2025
End of data collection, i.e., date from which the analytical dataset is completely available	Jul 2032
Annual Update Report 1	30 Oct 2025
Annual Update Report 2	30 Oct 2026
Annual Update Report 3	30 Oct 2027
Annual Update Report 4	30 Oct 2028
Annual Update Report 5/Interim report 1	30 Oct 2029
Annual Update Report 6	30 Oct 2030
Annual Update Report 7	30 Oct 2031
Registration in the EU PAS register	Withing 30 days after protocol endorsement by EMA
Final Extract	Jul 2032
Final report of study results	30 Oct 2033

Abbreviations: EMA, European Medicines Agency; EU, European Union; PAS, Post-authorization study

[1] Assuming EMA endorsement of the protocol 29 Sep 2022

5 Rationale and background

Multiple sclerosis (MS) is a chronic, inflammatory disease of the central nervous system that affects approximately 2.8 million people worldwide (Walton et al, 2020). The pathophysiology of multiple sclerosis involves B-cells and T cells (van Langelaar et al, 2020). Several disease modifying therapies (DMTs) which slow the progression of disease have been approved for the treatment of MS over the past decade (Hart and Bainbridge 2016). Better understanding of the crucial role played by the B-cells in the MS disease pathology has led to the advent of new DMTs that target B-cells resulting in the delay in progression of the disease. Anti-cluster of differentiation 20 (anti-CD20) monoclonal antibodies that target the CD20 receptors on B-cells, thereby affecting the B-cell-mediated cellular immunity, are one of the most recent classes of DMTs approved for treatment of relapsing and progressive forms of MS (Hart and Bainbridge 2016).

Kesimpta (ofatumumab) is a fully human monoclonal antibody that binds to a region of CD20, including the smaller and the larger loop of CD20 receptors.

In August 2020, United States (US) Food and Drug Administration (FDA) approved Kesimpta for the treatment of relapsing forms of MS including clinically isolated syndrome (CIS),

relapsing-remitting MS (RRMS) and active secondary progressive MS (SPMS) in adults (Kesimpta US Prescribing Information). Furthermore, in March 2021, European Commission granted the marketing authorization to Kesimpta for the treatment of relapsing forms of multiple sclerosis (RMS) in adults with active disease defined by clinical or imaging features (Kesimpta European Union [EU] Summary of Product Characteristics). To date, Kesimpta is the first and only self-administered, targeted B-cell therapy for the treatment of RMS, to include CIS, relapsing-remitting disease, and active secondary progressive disease, in adults. Kesimpta is the first B-cell therapy that can be self-administered once monthly at home via the Sensoready autoinjector pen (Bar-Or et al, 2021).

Observations during clinical trials suggest that Kesimpta may have the potential for an increased risk of infections, including serious bacterial, fungal, and new or reactivated viral infections. Furthermore, the long-term effects of exposure to Kesimpta on the risk of malignancies is not known.

In phase 3 trials of ofatumumab using teriflunomide as active comparator in MS patients, serious adverse events (SAEs) were reported in below 10% of ofatumumab recipients (Hauser et al., 2020) and the most commonly reported SAEs were among system organ class of 'infections and infestations', 'injury, poisoning and procedural complications', 'psychiatric disorders' and 'neoplasms benign, malignant and unspecified (including cysts and polyps)' (Kesimpta: EPAR - Public assessment Report, 2021). A serious infection was reported in 2.5% of MS patients treated with ofatumumab (Hauser et al, 2020). The most common infections seen in the pivotal trials of Kesimpta were nasopharyngitis, upper respiratory infections, and urinary tract infections, as reported for other CD20 targeting monoclonal antibodies rituximab and ocrelizumab (Hauser et al 2020, Roach and Cross, 2021). Neoplasms including basal-cell carcinoma, malignant melanoma in situ, recurrent non-Hodgkin's lymphoma, and invasive breast carcinoma have been reported in 0.5% of MS patients receiving Kesimpta (Hauser et al, 2020).

It is expected that Kesimpta will be used for the long-term treatment of MS, therefore, an investigation on the effects of long-term exposure on the risk of serious infections and malignancies is warranted. Further, as requested by the Committee for Medicinal Products for Human Use, a post-authorization long-term safety study in MS patients treated with ofatumumab in real-world settings (Category 3 Post-Authorization Safety Study [PASS]) was requested to be included in the Risk Management Plan of Kesimpta as one of the additional pharmacovigilance activities. Therefore, a non-interventional, long-term, retrospective PASS utilizing real-world data from existing registries from multiple European countries and Australia will be conducted to assess the risk of malignancies and serious infections following Kesimpta treatment in patients with MS during the study period.

6 Research question and objectives

6.1 Research question

Is Kesimpta initiation at any time, dose and/or duration during a MS patient's treatment journey associated with an increased risk of malignancy and infections compared to other DMTs initiation?

6.2 Objectives

6.2.1 Primary objective

- In patients diagnosed with MS, to compare the risk of 1) malignancy (except non melanoma skin cancers [NMSC], overall and for pre-defined type) and 2) late-onset infections* (overall, by type and seriousness) between Kesimpta-initiators and other DMT-initiators irrespective of therapy discontinuation or switch and 3) acute-onset infections* (overall, by type and seriousness) between Kesimpta-initiators and other DMT-initiators while on therapy.
  *including opportunistic infections

Intercurrent events can influence the effect of an intervention on the outcome. The intercurrent event can be related to the intervention, outcome or be unrelated. The below clarifies the question of interest using the estimand framework and particularly provides clarity on the handling of intercurrent events (Gogtay and Ranganathan, 2021). The primary estimand attributes are as follows:

- **Population**: Patients diagnosed with MS
- **Exposure of interest**: Kesimpta or other DMT treatment initiation (at any time, and/or duration)
- **Outcome**:
  1) Malignancy (except NMSC overall, and pre-defined type),
  2) Late-onset infections (overall, by type and seriousness, including opportunistic infections),
  3) Acute-onset infections (overall, by type and seriousness, including opportunistic infections)
- **Index date and exposure period/follow-up**:
  - **Index date**: day of Kesimpta or other DMT treatment initiation after Kesimpta launch date in respective country
  - **Start of the follow-up**: follow-up will start on the index date
  - **End of the follow-up**
    - **[for malignancy and late-onset infections]**: follow-up will end at event occurrence, or when the patient is censored from the database (e.g., due to emigration, end of follow-up in the data, or analysis cut-off date) whichever occurs first.
      - **[for acute-onset infections]**: follow-up will end at event occurrence or end date of the treatment plus drug specific wash-out periods or when the patient is censored from the database (e.g., due to emigration, end of follow-up in the data, or analysis cut-off date) whichever occurs first.
- **Handling of intercurrent event(s)**:
  - **[for malignancy and late-onset infections]**: A treatment policy strategy will be applied, and the event counted irrespective of treatment discontinuation or switch

- **[for acute-onset infections]**: A while on treatment policy strategy will be applied, and the event counted until treatment discontinuation plus drug specific wash-out periods.

- **Summary measure**: Causal contrast between Kesimpta-initiators vs. other DMTs initiators using hazard ratio (HR)

6.2.2 Secondary objectives

Characterize

- To **characterize** Kesimpta and other DMT-initiator populations and Kesimpta and other DMT use including but not limited to patient demographics, MS disease characteristics, duration of the Kesimpta and other DMT use, number of treatment switches.

Estimate

- To estimate, in patients diagnosed with MS, the incidence rate of **malignancies (overall except NMSC and for pre-defined type)** following Kesimpta and other DMT initiation irrespective of therapy discontinuation or switch.
- To estimate, in patients diagnosed with MS, the incidence rate of late-onset **infections (overall and by type)** following Kesimpta and other DMT initiation irrespective of therapy discontinuation or switch.
- To estimate, in patients diagnosed with MS, the incidence rate of **acute-onset infections (overall and by type)** following Kesimpta and other DMT initiation while exposed to Kesimpta or other DMT.
- To assess long term safety, estimate, in patients diagnosed with MS, the incidence rate of **SAEs (overall and by type)** following Kesimpta and other DMT initiation while exposed to Kesimpta or other DMT.

The following objectives will be explored **if feasible** in the data sources.

- To estimate, in patients diagnosed with MS, the incidence of **suicidal ideation, intestinal or bowel obstruction and sarcoidosis** following Kesimpta and other DMT initiation while exposed to Kesimpta or other DMT.

Compare

The following objectives will be explored if the observed number of cases allow for a statistical comparison.

- To assess long term safety, compare, in patients diagnosed with MS, the incidence of **SAEs (overall and by type)** between Kesimpta-initiators and other DMT-initiators while on therapy.

7 Research methods

7.1 Study design

This PASS is an observational, comparative, retrospective new user cohort study using longitudinal secondary data from national and multi-national real-world databases. This is a master protocol for conducting an observational long-term safety study. The master protocol

will be localized to the participating study countries and to the MSBase data source, to incorporate data specificities (e.g., availability) and impact on study design. Section 7.4 provides details on the selected data sources, as per a feasibility assessment performed in Q4-2021. As per study milestones, the reports will be provided annually starting 2025 (to allow accrual of patients in the data sources).

The target is to follow, for a minimum of 5 years, a minimum of 3000 Kesimpta-initiators and to compare these to other DMT-initiators during the same period. The study duration is planned for 8 years post Kesimpta launch in the participating study countries including a 3 year-accrual period. This allows for a theoretical minimum 5-year follow-up per patient, which is considered sufficient to evaluate malignancies and the long-term safety outcomes of this study. Note that, if needed, the accrual period may be lengthened to achieve the targeted sample size.

Due to the long-term nature of the study and the unknown Kesimpta use in real-world, the first annual update reports will be descriptive focusing on characterization of the Kesimpta-initiator population cohort and ensuring that the assumption on overlaps between cohort characteristics is reasonable. This would allow to inform any analytical adjustment to the planned analyses, via possible protocol amendments, if required. In annual reports, incidence and event rates of the outcomes of interest will be reported.

Infections will be classified in two groups, namely acute-onset and late-onset infections, depending on the latency/induction period from the end of treatment. The term “late-onset infections” will refer to expected late-onset infections including PML and HBV and varicella zoster virus (VZV) reactivation. The term “acute-onset infections” will refer to infections excluding previously qualified late-onset infections.

Outcomes of interest will further be categorised based on specific exposure windows and analytical conventions. Malignancies, late-onset infections and sarcoidosis will be handled as “Long-term outcomes”; while acute-onset infections, suicidal ideation and intestinal or bowel obstruction will be handled as “Short-term outcomes”.

For long-term outcomes, a treatment policy approach is considered and no restriction on the time horizon will be set i.e. the entire time from first drug intake until the end of the analysis cut-off date or censoring will be considered as time-at-risk. For short-term outcomes, which are expected to occur immediately after drug exposure, a while on treatment policy approach is applied and the time-at-risk will start at index date and extend up to end of wash-out period or risk windows accounting for the pharmacokinetic or pharmadynamic effect, as appropriate, for the considered DMT (Section 7.3.1 and Table 12-3).

When feasible, late-onset and acute-onset infections will be assessed for seriousness. Infections fulfilling the following criteria will be considered as serious: occur in close proximity to date of death, are life-threatening, requires in-patient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, or are otherwise considered as medically important. The serious infections definition may further be adapted in the localized protocols based on the data availability and reporting methodology in the country-specific data sources.

Combinations of databases from Denmark, Sweden, France, Italy and the MSBase data source were identified in the feasibility assessment to be suitable to address the questions of interest.

After Kesimpta launch date, any initiation of Kesimpta or any initiation of another DMT, not previously taken by the patient in his/her MS treatment history, will trigger a new index date and exposure period. Thus, a patient switching DMT regimen multiple times may contribute to both exposure cohorts and/or several times to the other DMT-initiator cohort. For the other DMT-initiator cohort, a new contribution to the cohort is considered only if the patient has no prior Kesimpta use.

Patients in both the Kesimpta and other DMT-initiator cohorts will be followed from study entry (index date) until the end of the study period (Section 7.2.2), study outcome of interest, end of data availability (emigration, end of follow-up in the data, or analysis cut-off date), or death, whichever occurs first.

Time to event of interest will be compared between the Kesimpta and other DMT-initiator cohorts for interim and final reports using Cox regression. The analyses will be conducted separately in each data source. Country-level effect size estimates will be pooled using suitable meta-analysis methods to compute overall estimates.

Using a new user design, i.e. considering treatment initiators in both the Kesimpta and the other DMT-initiator cohorts, will align patients in both treatment groups at a uniform point in time to start of follow-up (i.e., treatment initiation) and thereby reduce the potential for confounding from time-varying risk factors that may influence the decision to initiate treatment (Lund et al 2015). Considering an active comparator, i.e., using other DMT as the comparator cohort, will further reduce the potential for several types of bias, including confounding by indication (Lund et al 2015). The use of propensity score (PS) methods will allow to correct any possible imbalance in known clinical history between the two cohorts.

7.2 Setting and study population

7.2.1 Study population

The study will include adult patients with MS, aged 18 years or older who initiated Kesimpta or another DMT during the accrual period (Section 7.2.2). The study will utilize longitudinal secondary data from large healthcare data sources in planned participating countries/databases: Denmark, Sweden, France, Italy, and MSBase. An overview of data sources by country is shown in Table 7-2 and described in detail in Annex 4.

7.2.2 Study time frames

An overview of study times frames is presented in Figure 7-1.

The **study period** is country-specific and defined to start at Kesimpta launch date in the respective countries. The end of the study period will be country-specific and will correspond to the date of data extraction (last available data for analysis i.e., analysis cut-off date) in each country for individual study reports mentioned in the milestones (Section 4)

The **accrual period** will be the period in which the patients will be identified i.e., initial inclusion/exclusion criteria assessed. The study accrual period will be country-specific and start at Kesimpta launch date in the respective country and is planned to last for 3 years.

The **initial index date** refers to the first index date qualifying the patient for study entry.

For identified patients, the **index date(s)** will be defined as the date of Kesimpta and/or other DMT initiation. After Kesimpta launch date, any initiation of Kesimpta or another DMT, not previously taken by the patient in his/her MS treatment history, will trigger a new index date and exposure period. Thus, a patient switching DMT regimen multiple times may contribute to both exposure cohorts and/or several times to the other DMT-initiator cohort with several index dates. For the other DMT-initiator cohort, a new contribution to the cohort is only considered if the patient had no prior Kesimpta use.

The **pre-index** will be defined as a time period ending on the day before the index date.

The **follow-up period** will be defined as the time period between the index date and the end of the study period (analysis cut-off date) or censoring, whichever comes first (Section 7.3.3). Because patient accrual is planned to end 3 years after Kesimpta launch date, all patients will have a potential follow-up of minimum 5 years for long-term outcomes.

7.2.3 Selection criteria

Patients fulfilling all the inclusion and none of the exclusion criteria will be included in the study. Definitions of the criteria related to medication (for study exposure and covariates) and diagnoses of diseases (for study outcomes and covariates) are provided in Annex 3.

7.2.3.1 Inclusion criteria

Study level

Patients must fulfill all of the following criteria to be **included in the study** (Table 12-1):

- Aged 18 years or older at the index date
- A diagnosis of MS before/at the index date
- Kesimpta or other DMTs initiation within 3 years of Kesimpta launch date in respective database countries

Cohort level

Patients must fulfill cohort-specific inclusion criteria to be included in the respective cohort during the study period:

- **Kesimpta cohort:** Kesimpta initiation with no prior Kesimpta exposure history
- **Other DMT-initiator cohort:** initiation of a new DMT not previously recorded in-patient's prior MS treatment history.

7.2.3.2 Exclusion criteria

Study level

Patients fulfilling any of the following exclusion criteria will be excluded from the study (Table 12-2):

- Previous exposure to Kesimpta or specific other DMT before/at the initial index date

Cohort level

- **Kesimpta cohort:** prior Kesimpta exposure

- **Other DMT-initiator cohort:** prior exposure to Kesimpta or to the same DMT.

The included DMTs, corresponding Anatomical Therapeutic Chemical (ATC) codes, and risk windows are presented on Table 12-3.

7.3 Variables

The variables used in this study are divided into exposure variables (Section 7.3.1), outcome variables (Section 7.3.2), censoring variables (Section 7.3.3) and other variables (Section 7.3.4). The availability and content of specific variables are expected to vary between data sources, e.g., according to local coding system. This master protocol provides common variable definitions which will be adapted to the local definition of the variables using local specifications and/or coding systems in each data source. Specific variable definitions will be detailed and provided as part of local protocols and the statistical analysis plan (SAP).

Table 7-1 Summary of the variable of the study

	Variables	Defined in Section
Inclusion and exclusion criteria	Inclusion criteria, study level: • Aged 18 years or older at the index date • Diagnosis of MS before/at the index date • Kesimpta or other DMTs initiation within 3 years of Kesimpta launch date in respective data base countries Cohort level: • Kesimpta cohort: Kesimpta initiation with no prior Kesimpta exposure history • Other DMT-initiator cohort: initiation of a new DMT not previously recorded in-patient's prior MS treatment history Exclusion criteria, study level: • Previous exposure to Kesimpta or other DMT before/at the index date Cohort level: • Kesimpta cohort: prior Kesimpta exposure • Other DMT-initiator cohort: prior exposure to Kesimpta or to the same DMT	7.2.3.1 7.2.3.2
Exposure assessment	Exposure cohorts: • Kesimpta-initiator • Other DMT-initiator	7.3.1
Outcomes	Primary: • Overall malignancies excluding NMSC • Pre-defined malignancies • Late-onset infections (overall, by type and seriousness, including opportunistic infections)	7.3.2

	Variables	Defined in Section
	• Acute-onset infections (overall, by type and seriousness, including opportunistic infections) Secondary: • Serious adverse events (overall, by type, if feasible) • Other: Suicidal ideation, intestinal or bowel obstruction and sarcoidosis (if feasible)	
Censoring variables	• Death • End of data availability	7.3.3
Other variables	Characteristics related to demographics, lifestyle, MS disease, Medical history, comorbidities and comedications	7.3.4

Abbreviations: DMT, Disease modifying therapy; MS, multiple sclerosis; NMSC, non-melanoma skin cancer

7.3.1 Exposure assessment

7.3.1.1 Study drugs

Exposure to Kesimpta or other DMT-initiator cohort will be ascertained from recordings of prescriptions, prescriptions dispensed at community pharmacies, or administration in the hospital setting as available in the different data sources. The master protocol will be localized to the participating study countries and information on study drug availability will be assessed.

Any new initiation of drugs presented in Annex 3, Table 12-3 (without prior Kesimpta exposure) will qualify the patient for inclusion in the other DMT-initiator cohort. If any new drugs are introduced in any of the study countries during the study period, these would also be considered for the comparator group.

7.3.1.2 Definition of exposure period of malignancies, late-onset infections, and sarcoidosis

For malignancies, late-onset infections, and sarcoidosis, as primary approach, a treatment policy strategy will be applied, and the event counted irrespective of treatment discontinuation or switch.

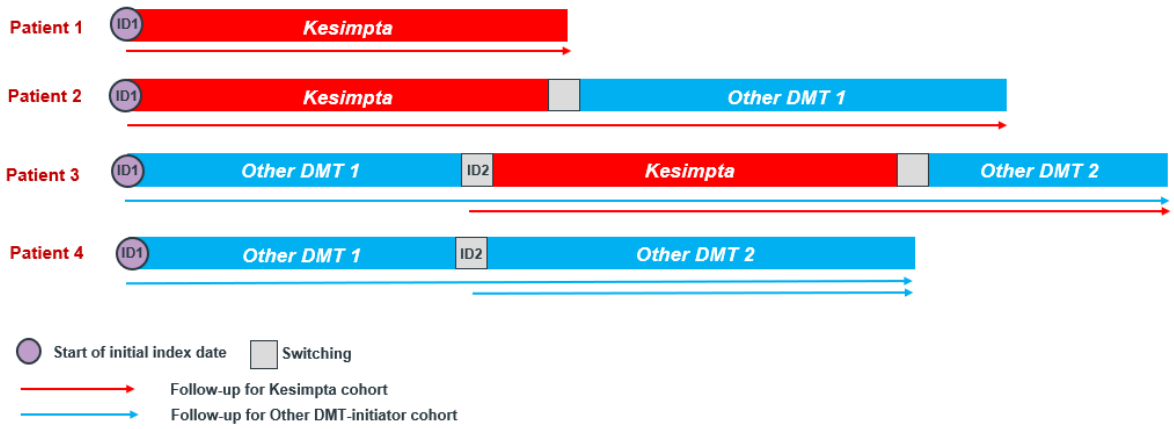
Therefore, each new Kesimpta or DMT initiation triggers a new index date, if cohort-specific inclusion criteria is fulfilled. The exposure period for malignancy, late-onset infections and sarcoidosis will start at the initial index date (the first date when the prescription was dispensed/issued/administered depending on available information) and end at the end of the analysis cut-off date, outcome or censoring, whichever comes first (Section 7.3.3).

In sensitivity analyses, accounting for possible latency period and recognizing that these remain unknown, arbitrary 6 and 12-month latency periods will be considered respectively. These alternative definitions will only apply to patients with more than 6 (or 12) months of follow-up.

In that case, the exposure period will start at the initial index date (i.e., the first date when the prescription was dispensed/issued/administered depending on available information) + 6 (or 12) months and end at the analysis cut-off date or censoring, whichever comes first (Section 7.3.3).

The primary exposure definition is illustrated in Figure 7-1.

Figure 7-1 Cohort and exposure period attribution for malignancies, late-onset infections, and sarcoidosis



Abbreviations: DMT, Disease modifying therapy; ID, index date

Note: The Patient 2 is not assigned ID2 and Patient 3 is not assigned ID3 since cohort level inclusion and exclusion criteria are no longer fulfilled at that moment in time for these patients. Patient 4 is assigned ID2 due to initiation of another DMT drug different from the first and no previous Kesimpta use (Section 7.2.3)

7.3.1.3 Definition of exposure period for acute-onset infections, SAEs, suicidal ideation, intestinal or bowel obstruction

In the primary definition of exposure period for acute-onset infections, SAEs and suicidal ideation, intestinal or bowel obstruction, a "while on treatment" policy strategy will be applied to handle treatment switching and discontinuation. A patient will be considered at-risk from the start of the index drug prescription/ dispensation/ administration (depending on available information, Annex 3, Table 12-3) until the end date of the treatment plus drug specific wash-out periods (risk window as per Table 12-3), event or censoring date (e.g., due to emigration, analysis cut-off date), whichever comes first (see Section 7.3.3).

If an event occurs in overlapping at-risk periods (i.e., after the initiation of a new DMT but still within the previous DMT wash-out period), both DMTs will be credited the time and event.

Conditional on inclusion/exclusion criteria been fulfilled, a patient with multiple index dates will contribute with multiple at-risk periods to the study.

As secondary approach, while on treatment analysis without risk window may be considered.

The exposure definition for these outcomes will be based on the pharmacokinetic effect rather than based on the pharmacodynamic effect for Kesimpta. The exposure definition for Kesimpta for these outcomes will be considered as the end date of the treatment plus 5 times the drug half-life (i.e. 85 days). A patient will be considered at-risk from the start of the index drug

prescription/ dispensation/ administration (depending on available information, Table 12-3) until the end date of the treatment plus 85 days for Kesimpta or drug specific wash-out periods (risk window as per Table 12-3), event or censoring date (e.g., due to emigration, analysis cut-off date), whichever comes first (see Section 7.3.3).

7.3.2 Outcomes

As the availability of data related to certain infections may vary among the databases and over time, detailed operational definitions for the identification of infections in different disease coding systems used by data sources will be defined in the SAP.

Outcomes will be classified in two categories requiring distinct analytical conventions. Malignancies and late-onset infections and sarcoidosis will be handled as "long-term outcomes"; while acute-onset infections, SAEs, suicidal ideation and intestinal or bowel obstruction will be handled as "short-term outcomes".

Primary outcome

Overall malignancies (excluding NMSC): Any occurrence of malignant neoplasms during the follow-up period, excluding NMSC. If a patient is diagnosed with two or more malignancies during the follow-up, the follow-up to the first event will be considered. Benign tumors will not be included.

Malignancies (pre-defined by type): Any occurrence of malignant neoplasms (pre-defined in Table 12-4) during the follow-up period. If a patient is diagnosed with two or more malignancies during the study period, the follow-up will be attributed to each malignancy when analyzed by malignancy type. Benign tumors will not be included.

Late-onset infections (overall, by type and seriousness, including opportunistic infections): All expected late-onset bacterial, viral, fungal, and parasitic infections, e.g., PML and HBV and VZV reactivation, as operationalized in Table 12-4, that occur during the study period will be included. A patient may contribute to more than one infection episode during follow-up. The seriousness criteria are taken as described in the "SAEs" definition provided below.

Acute-onset infections (overall, by type and seriousness, including opportunistic infections): All bacterial, viral, fungal, and parasitic infections, as operationalized in Table 12-4, that are expected to be have an acute-onset, excluding previously qualified late-onset infections, and occur during the exposure period will be recorded. A patient may contribute to more than one infection episode during follow-up. The seriousness criteria are taken as described in the "SAEs" definition provided below.

Secondary outcomes

Serious adverse events (overall and by type): Serious adverse events will be defined as events occurring in close proximity to date of death, are life-threatening, require in-patient hospitalization or prolongation of existing hospitalization, result in persistent or significant disability/incapacity, or are otherwise considered as medically important will be defined as SAEs. These will be classified as acute-onset events.

Note that the SAE definition may be adapted in the localized protocols based on the data availability and reporting methodology in the country-specific data sources.

Other outcomes: If feasible, suicidal ideation, intestinal or bowel obstruction and sarcoidosis, as defined in Table 12-4, may be assessed in the study.

7.3.3 Censoring variables during follow-up

The following events lead to censoring (Table 12-5):

- Death: date of death
- End of data availability (e.g., emigration, end of follow-up in the data, or analysis cut-off date (last available data for analysis in each country for individual study reports mentioned in the milestones [Section 4]))

7.3.4 Other variables

Variables potentially associated with malignancies and infections, such as demographic or lifestyle related variables including age, sex, body mass index, alcohol abuse and smoking, MS and medical history, and medication and exposure history, will be identified for study participants prior to the index date (Table 12-6), as available in the data sources.

Demographic characteristics:

- Sex
- Age at index date
- Geographic area
- Socio-economic status
- Calendar year of index date

Lifestyle characteristics, if available:

- Body mass index (height and weight)
- Alcohol/substance abuse
- Smoking

MS disease characteristics:

- MS disease duration at index date
- Duration of MS treatment at the index date
- MS type at index date (if available), e.g. RRMS, SPMS, Primary-Progressive MS (PPMS), Progressive-Relapsing MS (PRMS)
- Expanded Disability Status Scale (EDSS) or a proxy measure (if EDSS not available) prior to index date (using the measurement closest to index date) and during follow-up
- Number of relapses (including date of onset), including glucocorticoid treatment (yes/no) within one year prior to enrolment and during follow-up
- John Cunningham virus antibody status positive and negative at index date, if available

Medical history, health care utilization:

- Healthcare resource utilization

Medical history, other comorbid conditions:

- Cardiovascular disease
- Respiratory disease
- Gastrointestinal disease
- Metabolic disease
- Musculo-skeletal disease
- Mental disorders
- Autoimmune disease other than MS
- Immunodeficiencies
- Prior infection
- Previous malignancies
- Family history of malignancy, if available
- SARS-COV-2 infection

Medication and exposure journey:

- Kesimpta use: doses and all the starting and stopping dates, as indicated by relevant information on prescription or dispensation from specific data sources, major discontinuation reason, if available
- Other DMT use: generic or trade name, dose, as well as the starting and stopping dates, as indicated by relevant information on prescription or dispensation from specific data sources, major discontinuation reason, if available
- Concomitant medications: all non-DMT prescriptions including corticosteroids and prior immunosuppressive and immunomodulatory agents
- Vaccines: name (e.g., herpes zoster vaccine) and date, if available
- Radiation exposure, if available

Some of the baseline characteristics and covariates, e.g., demographic and lifestyle characteristics might be partly recorded in data sources as per findings of performed feasibility assessment. In addition, covariates related to MS characteristics, such as EDSS and discontinuation reasons, are exclusively recorded in MS quality registers. Thus, the final list of the covariates will be assessed in the localized protocols and baseline characteristics and covariates to be included in individual statistical analyses will be detailed in the SAP.

7.4 Data sources

All information on exposure and outcomes of the study population are derived from established electronically recorded longitudinal secondary data sources in Denmark, Sweden, France, and Italy, and in multi-country data source, MSBase. The use of these data sources will enable the inclusion of a large number of MS patients, including data on drug prescriptions, malignancy and infection outcome, medical history, and other clinical characteristics of MS patients.

Furthermore, the inclusion of eligible patients from multiple countries increases the generalizability of the results.

The planned data sources were selected for the study based on a feasibility assessment performed in Q4-2021. The selected data sources are presented in Table 7-2 together with size of the active MS population, Kesimpta launch date in respective country and the data lag of each data source. The following section provides a description of the selected data sources including information on data-source-related capabilities and limitations. Details on data sources evaluated and not selected including rationale for not selecting the data source are presented in Annex 4.

Table 7-2 Data sources selected for the study

Country	Size of active MS population	Kesimpta launch date	Data sources	Data lag
Denmark	~17,600	Jan-2022	Danish Civil Registration System (CPR)	Unknown[2]
			Danish Register of Medicinal Product Statistics (RMPS)	2 months
			Danish National Hospital Medication Register (Sygehusmedicinregisteret, SMR)	2 months
			Danish National Patient Register (NPR-Den)	2 months
			Register of Laboratory Results for Research (RLRR)	5 weeks
			Danish Cancer Register	13 months
			Danish Cause of Death Register	13 months
			The National Health Insurance Service Register (NHISR)	2-4 months
			The Danish COVID-19 Surveillance Database	Unknown[2]
			Danish Multiple Sclerosis Registry (DMSR)	1-6 months
Sweden	~21,000	Oct-2021	Total Population Register	1 month
			Swedish Prescribed Drug Register (SPDR)	2-4 months
			Swedish National Patient Register (NPR-Swe)	6-9 months
			Swedish Cancer Register	12 months
			Swedish Cause of Death Register	6 months
			Swedish Infectious Diseases Register (SmiNet)	Unknown[2]
			National Quality Registry for Neurological Care	Unknown[2]
France	~51,846	Sep-2021	Observatoire Français de la Sclérose en Plaques (OFSEP)	6-8 months
Italy	~26,470	Jul-2022 (expected)	Italian MS Registry (IMSR)	Months[2]

Country	Size of active MS population	Kesimpta launch date	Data sources	Data lag
MSBase[1]	~27,500	Q2-2021 to Q2-2022 (expected)	Australia, Belgium, Czech Republic, Netherlands, Spain, Turkey, UK	Months[2]

Abbreviations: MS, multiple sclerosis; UK, United Kingdom

[1] Pre-selected countries (Australia, Belgium, Czech Republic, Netherlands, Spain, Turkey, and UK) can be replaced with or be supplemented by data from other countries included in MSBase

[2] Will be explored further in localized protocol

7.4.1 Data sources and limitations

In the Nordics, the selected data sources are linked at country-level to form one database per country. Knowledge on the limitations of the databases in relation to this study may evolve with further investigations and once an operational connection is established with the data holder. A detailed description of the selected data sources is available in Annex 4.

7.4.1.1 Denmark

The Danish database is formed through linkage of several national Danish registries including, among other data sources, the Danish Civil Registration System, which provides all Danish residents with a unique personal identification number allowing for follow-up until death or emigration; the Danish Cause of Death Register, which contains information on the cause of death in ICD-10 (10th revision of the International Statistical Classification of Diseases and Related Health Problems) codes; the Danish Multiple Sclerosis Registry (DMSR), which contains data on all Danish residents diagnosed with MS including their treatment history and reported adverse treatment events; and the Danish Cancer Register, which contains patient-level data on incident cancers in the Danish population; and the Danish National Patient Register (NPR-Den), which has a national coverage of healthcare encounters and in- and outpatient care given in public hospitals with diagnoses recorded in ICD-10 codes. The Danish data sources have a coverage of the full Danish population and are considered high-quality.

Limitations of this database include that some information in the DMSR are only available as free text clinician notes. Danish Cancer Register and Danish Cause of Death Register have a lag time up to 13 months. Additionally, the Danish Hospital Patient Medication Register is not yet released for research, however, it is expected to be released during 2022. In Denmark, to protect patient's privacy, a rule of reporting small patient counts applies. Patient counts <3 or if protection is not enough, <5 will be masked.

7.4.1.2 Sweden

Multiple Swedish national registries are linked through unique personal identification numbers to form the Swedish database used in this study. These registries include, among other data sources, the Swedish National Patient Register (NPR-Swe), which has a national coverage of healthcare encounters and in- and outpatient care given in public hospitals with diagnoses recorded in ICD-10 codes; the Swedish Cause of Death Register, containing information on all deaths and their causes in Sweden with a long track-record of high completeness; the National Quality Registry for Neurological Care, which is the main source of information on MS disease

in our study; and the Swedish Cancer Register, which records newly discovered cancers in Swedish population. The data is considered of high-quality.

The Swedish database is limited by the lag time up to 12 months in Swedish Cancer Register. Data lag in Swedish Infectious Diseases Register (SmiNet) and National Quality Registry for Neurological Care is unknown and will be explored further in localized protocol. Additionally, to protect patient's privacy a rule of reporting small patient counts applies. Patient counts <5 will be masked.

7.4.1.3 France

The French Observatory of Multiple Sclerosis (OFSEP), established in 2003, is a national register actively covering data from 37 centers and representing about a quarter to a third of all MS patients in France. The register is limited by the fact that information on adverse events (AEs) is reported by the patient.

7.4.1.4 Italy

The Italian MS Registry (IMSR) includes data on approximately 60% of MS population in Italy. The register does not collect information on patients with severe MS, e.g., patients with high EDSS score and low mobility - EDSS >8.0. Additionally, until December 2019, comorbidities and AEs were collected as a free text, while from January 2020, ICD-9 and Medical Dictionary for Regulatory Activities (MedDRA) coding system were used respectively. Additionally, all AEs are self-reported by the patient. Detailed information on data lag will be explored further in localized protocol.

7.4.1.5 MSBase

MSBase is a multi-center, multi-country longitudinal observational MS database with 39 participating countries from several continents. In feasibility assessment performed in Q4-2021, pre-selected 7 MSBase countries Australia, Belgium, Czech Republic, Netherlands, Spain, Turkey, and United Kingdom were assessed in detail. However, these pre-selected countries can be replaced with or be supplemented by data from other countries included in MSBase, if needed. Its main limitations include the limited availability of data on patient disease history, medications for other diseases than MS and lifestyle factors, such as smoking and alcohol use. Patient comorbidities and concomitant medications other than MS treatments, e.g., female hormone use, are partially collected in the register. Most importantly, the information on SAEs, including infection diseases and malignancies, is recorded retrospectively and only if a patient shares this information with treating specialist. Detailed information on data lag will be explored further in localized protocol.

7.5 Study size/power calculation

The study is planned to follow-up, for a minimum of 5 years, a minimum of 3000 Kesimpta-initiators. These patients are planned to be accrued over a 3-year period across all databases.

The below calculations provide the expected effect size (expressed in terms of HR) that can be ruled out if the Kesimpta cohort size is 3000 or 5000 and Kesimpta:Comparator (K:C) ratio is

1:1, 1:3 or 1:10. Due to the observational nature of this study, neither the size of the Kesimpta cohort nor the K:C ratio can be initially known. Therefore, several scenarios are provided.

In addition, to illustrate the wide range of incidences of the outcomes under study, calculations were conducted for several outcomes of interest, such as any malignancy excluding NMSC, breast and cervical cancers (only for females), colorectal cancer, herpes and candida-related infections, upper respiratory tract infections and pneumonia, HBV and PML.

In the below calculations, the meta-analysis setting and possible switching of patients between treatment cohorts were not accounted for.

For a given Kesimpta sample size (n=3000 vs. n=5000), assuming a 3-year-accrual time, a minimum of 5-year follow-up per patient and a yearly loss to follow-up rate of 5% based on investigator experience with selected databases, the expected total time-at-risk can be estimated as follows:

$$E(\sum T_i) = \frac{n}{\lambda}\left[1 - \frac{\exp(-\lambda\tau)\{\exp(\lambda c)-1\}}{\lambda c}\right] \quad (1)$$

where T_i is the patient-time contributed by patient i, n is the total number of Kesimpta patients (e.g. 3000), λ is the drop-out rate (0.05 per year), c is the accrual period (3 years), and τ is the sum of accrual and follow-up periods (3 + 5 = 8 years).

Assuming there is no difference between Kesimpta and Comparator cohort and an exponential distribution for study outcome occurrence, the expected number of events for Kesimpta, d_k, is obtained by multiplying the total time-at-risk *(1)* with the outcome-specific incidence rate. Then, the log (HR) distribution can be approximated as follows (Scosyrev, Glimm 2017):

$$Log(HR) = N(0, \frac{(r+1)^2}{r*d}) \quad (2)$$

where r is the allocation ratio vs. Kesimpta (e.g. 3 when considering K:C ratio as 1:3) and d is the estimated number of events expected to be observed in total ($d=r*d_k$). This approximation holds for large samples if the hazard in each group is constant over time, the HR is close to one and the proportional hazards assumption is valid.

The HR and 95% CI upper limit can be derived based on the asymptotic distribution of log-HR *(2)*. With 80% power, the effect size (HR) that can be ruled out is determined as the 80th percentile of the 95% CI upper limit distribution. Table 7-3 provides the obtained estimates for 10'000 simulations. Calculations were done using R 4.1.2 (R Core Team, 2020).

Table 7-3 Effect size (HR) to be ruled out with 80% power by Kesimpta cohort size and Kesimpta:Comparator (K:C) ratio per incidence rate

Outcome	Incidence rate per 1000 PY in comparator	Expected number of events in Kesimpta cohort	Expected HR to be ruled out*		
			K:C ratio 1:1	K:C ratio 1:3	K:C ratio 1:10
Kesimpta n patients = 3000					
Malignancy[a]					
Any malignancy (excl. NMSC)	4.08	68	**1.61**	**1.48**	**1.43**
Breast cancer (female)**	1.52	19	2.48	2.10	**1.96**

Outcome	Incidence rate per 1000 PY in comparator	Expected number of events in Kesimpta cohort	Expected HR to be ruled out*		
			K:C ratio 1:1	K:C ratio 1:3	K:C ratio 1:10
Colorectal cancer	0.22	4	8.17	5.44	4.62
Cervical cancer (female)**	0.18	2	13.86	8.72	7.14
Infections					
PML[c]	0.04	1	-	-	-
Herpes-related infections[b]	20.05	333	**1.25**	**1.19**	**1.18**
Candida-related infections[b]	34.01	565	**1.18**	**1.15**	**1.13**
Upper respiratory tract infections[b]	85.41	1418	**1.11**	**1.09**	**1.08**
Pneumonia (and influenza)[b]	65.23	1083	**1.13**	**1.10**	**1.09**
HBV[b]	1.87	31	2.03	1.79	**1.69**
Kesimpta n patients = 5000					
Malignancy[a]					
Any malignancy (excl. NMSC)	4.08	113	**1.46**	**1.35**	**1.32**
Breast cancer (female)**	1.52	32	2.02	**1.78**	**1.69**
Colorectal cancer	0.22	6	4.97	3.72	3.26
Cervical cancer (female)**	0.18	5	7.78	5.38	4.54
Infections					
PML[c]	0.04	1	-		
Herpes-related infections[b]	20.05	555	**1.18**	**1.15**	**1.13**
Candida-related infections[b]	34.01	941	**1.14**	**1.11**	**1.10**
Upper respiratory tract infections[b]	85.41	2364	**1.08**	**1.07**	**1.06**
Pneumonia (and influenza)[b]	65.23	1806	**1.10**	**1.08**	**1.07**
HBV[b]	1.87	52	**1.74**	**1.56**	**1.50**

Abbreviations: HBV: hepatitis B virus; HR: hazard ratio; NMSC: non-melanoma skin cancer; PML: progressive multifocal leukoencephalopathy; PY: person-years

[a] Norgaard M, Veres K, Sellebjerg F, et al. Incidence of malignancy in multiple sclerosis: A cohort study in the Danish multiple sclerosis registry. Multiple Sclerosis Journal 2021

[b] Castelo-Branco A, Chiesa F, Conte S et al. Infections in patients with multiple sclerosis: A national cohort study in Sweden. Multiple Sclerosis and Related Disorders. 2020;45:102420

[c] Schwab N, Schneider-Hohendorf T, Melzer N, et al. Natalizumab-associated PML: challenges with incidence, resulting risk, and risk stratification. Neurology 2017;88(12):1197-1205

*obtained via simulations (10'000 simulations)

** assuming 75% of female in the population (Harbo et al, 2013)

7.6 Data management

This study will follow the relevant chapters of the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP).

The data management processes differ by country and data source. Security processes will be in place to ensure the safety of all systems and data. The data is stored in a secured restricted area and it cannot be accessed by anyone except authorized study staff. Appropriate data storage and archiving procedures will be followed. Standard procedures will be in place to restore files in the event of a hardware or software failure.

7.6.1 Data owners

Data will be extracted locally by each data owner and analyzed either by data owner or by [redacted] (see Section 7.6.3 and Section 7.6.4). Extraction of data and data management will be done according to registry-specific procedures. Routine procedures pre-specified and approved by each disease registry will include checking electronic files, maintaining security and data confidentiality, following analysis plans and performing quality control checks of all programs, when appliable. The datasets and analytic programs relevant for the execution of this study will be stored according to the registries' or [redacted] procedures to allow retrospective data or program interrogation and review as well as further analyses, if needed.

R [R Core Team (2020)] version 3.2 or later, SAS [SAS Institute., Cary, NC, USA] version 9.2 or later, or other appropriate statistical software will be utilized to access the raw data, to manage the analytic datasets, and to conduct data analysis.

[redacted]

7.6.3 Patient-level data transfer

Patient-level data transfer will be performed in countries where [redacted] will be the data processor. The countries will be specified in the localized protocol. After data have been transferred to [redacted] they will be stored in a secure server managed by [redacted]. Data management procedures and checks will be built to map the source data into a consolidated unified format. This approach will allow identification of any inconsistencies in the data received.

A unique dummy Study Identification Number (SID) will be created for each patient, prior to data delivery to [redacted]. The SIDs will be used for data linkage on patient-level. Thereby the researchers at [redacted] will only have access to data where individuals cannot be directly identified.

7.6.4 Aggregate-level data transfer

Aggregate-level data transfer will be performed in countries where the data source is the data processor. The countries will be specified in the localized protocol Once the data have been transferred to [redacted] they will be stored in a secure server managed by [redacted]. [redacted] will perform meta-analyses for the combined data.

Within each data source, a common data model will be used to generate aggregated datasets in a pre-defined standard format. The common data model defines appropriate level of aggregation and minimum information required to enable the statistical analysis of the study objectives and will be described in detail in the SAP. The contents of the common data model will be reviewed and approved by the relevant data sources before data extraction and transfer. [redacted] will provide the programming code to be executed on locally generated common data model. This will allow to generate standardized aggregated results and ensure that harmonized statistical methods are used. These aggregated results can then be meta-analyzed by [redacted] as described above.

Full details of data management will be specified in the data preparation plan and programming specification.

7.7 Data analysis

7.7.1 Setting

Since this study is run across databases, analyses will be performed in two stages (see Section 7.6). First, data will be analyzed locally or by [redacted] for each data source following a common statistical methodology (Section 7.6.1). This will be achieved using a locally adapted common data model and common statistical programs provided by [redacted] (the designated coordinator for the study). Second, the aggregated data or stratified summaries, as appropriate, from each data source will be provided to [redacted] to conduct integrated analyses using meta-analytical methods (Section 7.7.4.2).

7.7.2 General analytical conventions

For continuous variables, the following measures will be reported: the number of patients with non-missing data, the number of patients with missing data, mean, standard deviation, median, quartiles, and minimum and maximum, whenever possible.

For categorical variables, the number of non-missing observations (number and percentage of total non-missing; n, % of total), as well as the number of patients with missing data will be reported.

The total number of events and respective rates per 1000 person-years (PY) with 95% CIs will be provided. Two types of rates may be presented:

- **Event rates:** rates per 1000 PY, including all events reported within the qualifying exposure period (i.e. multiple events allowed)
- **Incidence rates:** rates per 1000 PY, using first events only (i.e. censoring at the time of first events)

7.7.3 Analysis for annual reports

Annual reports will be cumulative and descriptive in nature. Each data owner will describe patient characteristics, treatments, outcomes and follow-up accrued by the data cut-off date.

The descriptive analyses will be particularly important to better understand the population treated with Kesimpta and ensure that the assumptions made regarding the overlap in cohort characteristics are reasonable. Further, these analyses will inform whether any adjustments are warranted to the planned comparative analyses.

Exposure periods are defined as per Section 7.3.2.

7.7.3.1 Description of patient characteristics

The descriptive analysis will focus on the number and percentage of unique patients per cohort (overall and newly contributing by calendar year). Patients will be described in terms of demographic characteristics (e.g., age, gender), potential confounders for malignancy and infections, prior therapies received and comorbidities at index date.

Patients contributing to both study cohorts will be reported in both cohorts.

7.7.3.2 Description of treatments at initial index date and during follow-up

The number of patients exposed to several DMTs, contributing to several cohorts and/or several times to the comparator cohort will be described including the number of index DMTs per patient will be described.

The total and mean duration of follow-up irrespective of therapy change and accounting for therapy change, i.e., contributing time on Kesimpta or other DMTs (overall and per individual DMT) will be summarized along with switching rates. The frequency of DMT treatment sequences will be provided.

7.7.3.3 Outcome descriptive analysis

Frequencies and raw (unadjusted) incidence rates with 95% confidence intervals (CIs) (expressed in 1000 person-years) for the pre-specified outcomes in both cohorts will be provided using Poisson regression or a corresponding method for over-dispersed data.

Depending on the nature of the outcome (malignancy, late and acute-onset infections, SAEs , suicidal ideation, intestinal or bowel obstruction and sarcoidosis), the appropriate exposure period will be considered (see Section 7.3.1). Whenever relevant, a gender restriction will be applied (e.g., breast cancer, cervical cancer).

In addition, for acute-onset infections and SAEs,

- the event rates may be provided to account for the possible recurring nature of outcomes,
- Kaplan Meier curves or equivalent may be provided considering 6-month intervals.

Note that if any acute-onset infections occur in an overlapping exposure period (i.e., after the initiation of a new DMT but still within the previous DMT wash-out period), both DMTs will be credited the time and event.

7.7.4 Analysis for interim and final reports

The final and interim analyses (planned 7 years after protocol finalization) will include both descriptive and comparative analyses. Descriptive analyses are described in Section 7.7.3.1, Section 7.7.3.2, and Section 7.7.3.3.

Whenever the observed number of cases allows statistical comparison, each data owner will conduct comparative analysis and provide source specific treatment effect estimates (HRs). [redacted] will estimate the overall treatment effect (HR and 95% CI) using methods described in Section 7.7.4.2.

7.7.4.1 Comparative analysis

Incidence rates will be compared using unadjusted and adjusted HR. Survival analysis including Cox proportional-hazard model will be used to obtain the HR estimates and 95% CIs. These analyses will focus on time to first event accounting for the appropriate exposure window for each outcome (Section 7.3.1).

Note that each outcome will be evaluated in an independent manner i.e., no competing risk principle is applied across outcomes. No adjustment for multiplicity will be considered. The assumptions of the models will be checked.

Whenever possible, the model will be stratified by naïve DMT user vs. non-naïve DMT users.

To account for the non-randomized nature of the comparison and to correct for possible imbalance in pre-treatment covariates, PS weighting methods (accounting for but not limited to a pre-specified list of confounders or variables with standardized difference > 0.1 (Flury and Riedwyl 1986)) will be used to ensure cohort comparability at index date (Austin 2011).

Overlap in the PS distribution across cohorts will be assessed graphically. In addition, covariate distributions before and after weighting will be assessed providing the standardized mean differences of each covariate.

Confidence intervals of outcome models will be obtained using robust standard errors or bootstrapped to account for weighting and possible multiple contributions per patient.

Potential confounders are likely to include, but will not be limited to, age, sex, calendar time, disease duration prior to treatment initiation, proportion of disease duration spent treated with DMT, EDSS, comorbidities, prior drug exposure, concomitant drugs (e.g., other concomitant immunomodulators/suppressants), pre-baseline relapse activity, and (if appropriate) country.

More detailed analyses corresponding to each of the study objectives will be described in full detail in the SAP as a separate document.

7.7.4.2 Combining results across data sources

The interim and final data analysis will provide the overall treatment effect estimate (HR and 95% CI), integrating estimates obtained across data sources. These will be obtained using a random effects model and the Hartung-Knapp estimator (Hartung, 1999; Hartung and Knapp, 2001). Heterogeneity will be assessed using Cochran's Q and the I^2 statistic. Results will be graphically displayed using forest plots.

7.7.5 Subgroup analysis

If sufficient data are available, subgroup analyses may be conducted including for those above 55 years of age.

7.7.6 Handling of missing data

Methods commonly used in non-interventional studies (NIS) utilizing secondary data for handling missing data, e.g. multiple imputation methods, may be applied, if needed. In descriptive analyses, missing data will be described separately and not included in the denominator for the calculation of the percentage or rates. Full details on handling missing data will be described in the SAP.

7.7.7 Sensitivity analysis

7.7.7.1 Sensitivity analysis: Latency period for malignancy

To take disease latency into account for malignancy, the analysis for the primary outcome will be complemented by analyses accounting for empirical latency period of six months post-index date. In these analyses, follow-up time and outcomes occurring six months after the index date are not considered. Time to first event will be calculated taking as reference the initial index date +6 months. PS weighting will be conducted including covariate status at the new reference date. The true latency period remaining unknown for most cancers, further sensitivity analyses will be conducted accounting for a 12-months latency period.

7.7.7.2 Sensitivity analysis: Risk window for acute-onset infections

Drug specific risk windows, based on wash-out period required for the drugs to be eliminated from the body, are considered in the main analysis, hypothesizing that the possible increase in risk, if any, remains until the complete elimination of the drug. However, there is lack of evidence for this mechanistic pathway and therefore, a sensitivity analysis without considering a risk window after treatment discontinuation or switch will be conducted.

7.7.7.3 Sensitivity analysis: Exclusion criteria for malignancy

Patients with cancer can have higher risk for a subsequent second cancer (new cancer) or a recurrence, due to hereditary factors, treatments of the first cancer, immunosuppression, age, hormonal and environmental influences, genetic predisposition and infection etc. (Demoor-Goldschmidt 2019). A sensitivity analysis excluding patients with cancer (of any type) before initial index date will be conducted.

7.8 Quality control

This study will be conducted according to the rules of 'Good Pharmacoepidemiology Practice' (GPP) and the 'Guideline on good pharmacovigilance practices (GVP) – Module VIII (Rev 1)' EMA/813938/2011 Rev 1. All aspects of the study from protocol development to the reporting of the results are conducted within the work-framework of [redacted] Quality Management System and in accordance with the following manual, operating procedures and work instructions.

According to the policies and procedures above a Quality Control plan for the study will be developed and executed, which will include quality control on study methodology, SAP, programming, data management and analysis, and study report including study results and conclusions. Furthermore:

- The principle of the independence of Quality Control applies.
- [redacted] project management will ensure that individuals responsible for the execution of specific Quality Control steps have knowledge, capability and experience which are adequate for the task.
- The result of the execution of the individual steps of the Quality Control plan will be documented, and include the required corrective actions, if any.
- The execution of any required corrective action will be documented.
- The executed Quality Control plan will be subjected to a final review and approval for sufficiency and completeness by the [redacted] project management team.
- The project management will also ensure that [redacted] employees assigned to the project are trained on protocol and project-specific procedures, as per [redacted] procedure.

The study will be conducted as prescribed in this protocol. Revisions to the protocol should be approved by the principal investigator and the marketing authorization holder (MAH). All changes to the protocol shall be documented as protocol amendments and when necessary, such protocol amendments will be delivered to relevant ethics committees, health authorities and register holders.

The study protocol has been written following the Code of Conduct by the ENCePP (EMA, 2018). The protocol also follows the key elements of the Guideline for GPP by International Society for Pharmacoepidemiology.

[redacted] the principal investigator, the MAH and individuals acting on their behalf commit to adhere to the rules of the ENCePP Code of Conduct in their entirety.

For information on storage of records, archiving of the statistical programming will be performed to generate the results and possible audits. Due to the study type (observational study using secondary databases) on-site monitoring will not be performed.

7.9 Limitations of the research methods

Although all possible measures will be taken to ensure the quality and robustness of the data, there are several limitations inherent to the study design and data sources that should be acknowledged.

7.9.1 Selection bias

The secondary data sources used in this study are considered to be a good representation of the patient population in the respective countries. In Denmark and Sweden, the MS registries cover approximately 100% and 80% of the MS patient population in the respective countries. Similarly, in France and Italy, the registry coverage is relatively high. Therefore, considering the largely representative patient population available in the selected data bases and the inclusion/exclusion criteria of the study, selection bias is less likely to occur in this study. However, it should be noted that the patient/database selection is constrained by the availability

of specific data points and database characteristics (See Annex 4) applied during the feasibility assessment for selection of databases for this study. Further, the patient identification period is relatively short and corresponds to an early market launch period. There could be a chance that patients included during the early market launch period are different from those who could have been included during the later years post market launch. The magnitude and direction of this bias may be evaluated by comparing the demographic characteristics of Kesimpta patients identified in the first year vs. those in the third year. In addition, the selected internal comparator guarantee validity of the comparison made.

7.9.2 Information bias

As in any study based on secondary data, the accuracy of the analyses depends on the accuracy of the available information in the included data sources. To minimize the risk of including inaccurate information, all retrieved data will be reviewed for possible inconsistences or implausible information. Missing information might also vary between the data sources. When applicable, the percentage of missing information will be reported.

7.9.3 Validity of exposure assessment

Relying on prescription or dispensation data for exposure assessment assumes perfect treatment adherence. Non-dispense or non-use of dispensed prescription may lead to exposure misclassification. However, exposure misclassification is expected to be non-differential between the exposure cohorts and not influence the study results.

In order to minimize the exposure misclassification of MS DMTs, the MS quality registers will be included in this study, to retrieve information on hospital administered DMTs.

The estimation of the risk window of different DMTs is based on pharmacokinetic data from different sources, which may lack in consistency and reliability. Applying a definition of risk window for the specific drugs that was shorter or longer than the true duration would lead to exposure misclassification. The strength and direction of any resulting bias would depend on the amount of misclassification. Also, the definitions of the risk window are estimated based on normal metabolism and liver and kidney function. Some comorbidities and polypharmacy might affect the length of the individual risk window of DMTs, which could lead to exposure misclassification. To account for exposure misclassification due to incorrect exposure risk window, a sensitivity analysis without a risk window will be conducted.

7.9.4 Outcome validity

The validity of the outcomes within our study is predicated on the coding within the databases themselves. For alignment purposes, and to limit possible misclassification, this study will rely on validated case algorithms when available. If a published algorithm is not available, a formal chart review will be considered.

7.9.5 Confounding

Baseline characteristics, including use of other DMTs, will be accounted for using Inverse Probability of Treatment Weighting methods. However, post-index exposure to carcinogenic and immunotoxic medications other than the considered exposures could be a source of

confounding if there is a differential use of these drugs between the exposure cohorts. Exposure to known carcinogenic and immunotoxic medications during the post-index period (including dosage and timing) may, therefore, also be considered for further adjustment.

The dose and timing of exposure to DMTs may carry a risk of confounding since changes of dosage could occur after the index date. It is difficult to estimate the extent of this confounding risk as the change of dosage might vary systematically between exposure cohorts based on the adverse reactions (AR) of the drugs.

Not all confounding factors will be captured or available for adjustment during analysis.

Residual confounding due to unknown or imprecisely measured variables may remain after adjustments.

7.9.6 Study size

For 5000 patients initiating Kesimpta, the HR that can be ruled out with 80% power are below 1.5 for most of the outcomes of interest, including any malignancy (excl. NMSC).

For a lower number of Kesimpta-initiators, however, it is possible that a substantial increase in risk of some safety outcomes may remain undetected.

7.10 Other aspects

Not applicable.

8 Protection of human subjects

This study was designed and shall be implemented and reported in accordance with the Guidelines for GPP of the International Society for Pharmacoepidemiology (ISPE 2016), the Strengthening the Reporting of Observational Studies in Epidemiology guidelines (Vandenbroucke et al, 2007), and with the ethical principles laid down in the Declaration of Helsinki.

This study is fulfilling the criteria of a ENCePP study' and follows the 'ENCePP Code of Conduct' (EMA 2016).

The localized protocols will be submitted and reviewed by the local Institutional Review Boards and/or Ethics Committees before start of study when required by the national guidelines. The study will not start in a country before written confirmation of a favorable approval from the Review Board and/or Ethics Committee has been received.

Any amendments to the protocol will also be submitted to the concerned Review Boards and/or Ethics Committee before implementation in the event of substantial changes.

9 Management and reporting of adverse events/adverse reactions

As this is a study based on secondary use of data, safety monitoring and safety reporting, where there is a safety relevant result, will be provided on an aggregate-level only; no reporting on an individual case level to Novartis is required.

In studies based on secondary use of data with a safety relevant result, reports of AEs/ARs will be summarized in the study report, i.e. the overall association between an exposure and an outcome will be presented. Relevant findings from the study report will be included in the periodic aggregated regulatory reports submitted to Health Authorities.

10 Plans of disseminating and communicating study results

Upon study completion and finalization of the study report, the results of this non-interventional study may be either submitted for publication and/or posted in a publicly accessible database of results. Publications will comply with internal Novartis standards and the International Committee of Medical Journal Editors (ICMJE) guidelines.

For applicable non-interventional PASS (in the EU or mandated by an EU Health Authority outside the EU), the final manuscript will be submitted to EMA and the competent authorities of the Member States in which the product is authorized within two weeks after first acceptance for publication.

11 References (available upon request)

Andrade SE, Shinde M, Moore Simas TA et al (2021). Validation of an ICD‐10‐based algorithm to identify stillbirth in the Sentinel System. Pharmacoepidemiol Drug Saf; 30:1175–1183.

Austin PC (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behav Res; 46(3):399-424.

Bar-Or A, Wiendl H, Montalban, X et al (2021). Rapid and sustained B-cell depletion with subcutaneous ofatumumab in relapsing multiple sclerosis: APLIOS, a randomized phase-2 study. Multiple Sclerosis Journal; 135245852110444. Available From: https://doi.org/10.1177/13524585211044479 (Accessed 11 February 2022).

Berkovich R, Fox E, Okai A et al (2021). Identifying disability level in multiple sclerosis patients in a U.S.-based health plan claims database. J Med Econ; 24(1): 46-53.

Demoor-Goldschmidt C, de Vathaire F. (2019). Review of risk factors of secondary cancers among cancer survivors. The British journal of radiology; 92(1093):20180390.

European Medicines Agency. Kesimpta, INN-ofatumumab: Summary Of Product Characteristics. Available from: https://www.ema.europa.eu/en/documents/product-information/kesimpta-epar-product-information_en.pdf (Accessed 11 February 2022).

European Medicines Agency (2016). The ENCePP Code of Conduct – for Scientific Independence and Transparency in the Conduct of Pharmacoepidemiological and Pharmacovigilance Studies. Available from: http://www.encepp.eu/code_of_conduct/documents/ENCePPCodeofConduct_Rev3amend.pdf (Accessed 11 February 2022).

European Medicines Agency (2018). The ENCePP Code of Conduct, Final report. EMA/721895/2018. Available from: https://www.encepp.eu/encepp/openAttachment/studyResultLatest/26950 (Accessed 11 February 2022).

European Medicines Agency (2021). Assessment report: kesimpta. Available from: https://www.ema.europa.eu/en/documents/assessment-report/kesimpta-epar-public-assessment-report_en.pdf (Accessed 11 February 2022).

Flury B K, Riedwyl H. (1986). Standard distance in univariate and multivariate analysis. The American Statistician; 40(3):249-251.

Gogtay NJ, Ranganathan P, Aggarwal R (2021). Understanding estimands. Perspect Clin Res; 12(2):106-112.

Harbo HF, Gold R, Tintoré M (2013). Sex and gender issues in multiple sclerosis. Ther Adv Neurol Disord; 6(4): 237-248. Hart FM, Bainbridge J (2016). Current and emerging treatment of multiple sclerosis. Am J Manag Care; 22:s159-70.

Hartung J (1999). An alternative method for meta‐analysis. Biometrical Journal: Journal of Mathematical Methods in Biosciences; 41(8):901-916.

Hartung J, Knapp G (2001). A refined method for the meta-analysis of controlled clinical trials with binary outcome. Stat Med; 20(24):3875-89.

Hauser SL, Bar-Or A, Cohen JA et al (2020). Ofatumumab versus Teriflunomide in Multiple Sclerosis. New England Journal of Medicine; 383(6):546–557.

Hernán MÁ, Brumback B, Robins, J M (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. Epidemiology: 561-570

Howard J (2014). Minimum latency & types or categories of cancer. Available from: https://www.cdc.gov/wtc/pdfs/policies/wtchpminlatcancer2014-11-07-508.pdf (Accessed 11 February 2022)

ISPE (2016). Guidelines for good pharmacoepidemiology practice (GPP). Pharmacoepidemiol Drug Saf; 25:2–10.

Lund JL, Richardson DB, Stürmer T (2015). The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. Curr Epidemiol Rep; 2:221–8.

Mason K, Thygesen, LC, Stenager E et al (2012). Evaluating the use and limitations of the Danish National Patient Register in register-based research using an example of multiple sclerosis. Acta Neurologica Scandinavica; 125:213–217.

Nørgaard M, Veres K, Sellebjerg F T et al (2021). Incidence of malignancy in multiple sclerosis: A cohort study in the Danish Multiple Sclerosis Registry. Multiple Sclerosis Journal–Experimental, Translational and Clinical, 7(4), p.20552173211053939.

R Core Team (2020). A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: https://www.R-project.org/ (Accessed 11 February 2022)

Roach CA, Cross AH (2021). Anti-CD20 B Cell Treatment for Relapsing Multiple Sclerosis.Frontiers in Neurology. Available from: https://www.frontiersin.org/article/10.3389/fneur.2020.595547 (Accessed 11 February 2022)

Schmidt M, Schmidt SA, Sandegaard JL et al (2015). The Danish National Patient Registry: a review of content, data quality, and research potential. Clin Epidemiol; 7:449-90.

Scosyrev E, Glimm E (2019). Power analysis for multivariable Cox regression models. Statistics in medicine; 38(1):88-99.

van Langelaar J, Rijvers L, Smolders J et al (2020). B and T Cells Driving Multiple Sclerosis: Identity, Mechanisms and Potential Triggers. Frontiers in Immunology; 11:760.

Vandenbroucke JP, von Elm E, Altman DG, et al (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Ann Intern Med; 147(8):W163-94.

Walton C, King R, Rechtman L et al (2020). Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition. Multiple Sclerosis Journal; 26(14):1816–1821.

12 Annexes

12.1 Annex 1 – List of stand-alone documents

None.

12.2 Annex 2 – ENCePP checklist for study protocols

Study title:
Kesimpta long-term retrospective safety study utilizing real-world data from existing multiple sclerosis registries and databases from multiple countries

EU PAS Register® number:
Study reference number (if applicable): OMB157G/Kesimpta/COMB157G2406

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				4
1.1.1 Start of data collection[1]	☒	☐	☐	4
1.1.2 End of data collection[2]	☒	☐	☐	4
1.1.3 Progress report(s)	☒	☐	☐	4
1.1.4 Interim report(s)	☒	☐	☐	4
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	4
1.1.6 Final report of study results.	☒	☐	☐	4

Comments:

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1 Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	6
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)	☒	☐	☐	5,6
2.1.2 The objective(s) of the study?	☒	☐	☐	6
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	6
2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☐	☒	
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

[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)	☒	☐	☐	7.1
3.2 Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	7.1
3.3 Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	7.1
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))	☒	☐	☐	6; 7.7.3.3.
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)	☒	☐	☐	9

Comments:

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.1 Is the source population described?	☒	☐	☐	7.2.1
4.2 Is the planned study population defined in terms of:				
4.2.1 Study time period	☒	☐	☐	7.2.2; 7.4
4.2.2 Age and sex	☒	☐	☐	7.2.1
4.2.3 Country of origin	☒	☐	☐	7.2.1; 7.4
4.2.4 Disease/indication	☒	☐	☐	7.2.1
4.2.5 Duration of follow-up	☒	☐	☐	7.2.2
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	7.2.3

Comments:

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)	☒	☐	☐	7.3.1

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☐	☐	☒	
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?	☒	☐	☐	7.1; 7.3.1.3.
5.6 Is (are) (an) appropriate comparator(s) identified?	☒	☐	☐	7.3.1.1

Comments:

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?	☒	☐	☐	6, 7.3.2
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	7.3.2
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)	☐	☐	☒	
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:

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☒	☐	☐	7.9.5
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☒	☐	☐	7.9.1

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	7.9.2, 7.9.3, 7.9.4, 7.9.5

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)	☐	☒	☐	

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)	☒	☐	☐	7.4 Annex 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)	☒	☐	☐	7.4 Annex 4
9.1.3 Covariates and other characteristics?	☒	☐	☐	7.4 Annex 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)	☒	☐	☐	7.4 Annex 4
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	7.4 Annex 4
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	7.4 Annex 4
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	Annex 3
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	Annex 3

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.3.3 Covariates and other characteristics?	☒	☐	☐	Annex 3
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	7.6.3

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	7.7
10.2 Is study size and/or statistical precision estimated?	☒	☐	☐	7.5
10.3 Are descriptive analyses included?	☒	☐	☐	7.7
10.4 Are stratified analyses included?	☒	☐	☐	7.7
10.5 Does the plan describe methods for analytic control of confounding?	☒	☐	☐	7.7
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☐	☒	☐	
10.7 Does the plan describe methods for handling missing data?	☒	☐	☐	7.7.6
10.8 Are relevant sensitivity analyses described?	☒	☐	☐	7.7.7

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)	☒	☐	☐	7.6
11.2 Are methods of quality assurance described?	☒	☐	☐	7.8
11.3 Is there a system in place for independent review of study results?	☒	☐	☐	7.8

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?	☒	☐	☐	7.9.1

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1.2 Information bias?	☒	☐	☐	7.9.2
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).	☒	☐	☐	7.9.5
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)	☒	☐	☐	7.9.6

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?	☒	☐	☐	8
13.2 Has any outcome of an ethical review procedure been addressed?	☐	☒	☐	
13.3 Have data protection requirements been described?	☒	☐	☐	8

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?	☒	☐	☐	3

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)?	☒	☐	☐	10
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	10

Comments:

Name of the main author of the protocol: [redacted]

Date: 23/May/2022

Signature: _____

12.3 Annex 3 – Definitions of inclusion/exclusion criteria, study drugs, outcomes, and other variables based on medication, diagnoses and procedures

The list of corresponding ICD-10 and MedDRA, if any, will be updated and adapted to local data source specific protocols as required. Additional diagnosis codes, surgical and procedure codes, and laboratory tests may be added to the definition of individual variables, depending on the availability in the data sources; exact definitions will be included in the data source specific local protocol.

Table 12-1 Inclusion criteria

Variable	ICD-10 code	Definition	Comments
Aged 18 years or older at the index date Age[1]	-	-	-
A diagnosis of MS (before/at the index date)[1]	G35.x	N/A	For data sources in which information on diagnoses of MS may be incomplete or missing, proxy ascertainment will be used to define this criterion (Berkovich et al., 2021)
Kesimpta or other DMTs initiation within 3 years of Kesimpta launch date in respective database country	-	-	-
For Kesimpta cohort: Kesimpta initiation with no prior Kesimpta exposure history	-	-	See Table 12-3 for ATC
For other DMT-initiator cohort: initiation of a new DMT not previously recorded in-patient's prior MS treatment history	-	-	See Table 12-3 for ATC

Abbreviations: ATC, Anatomical Therapeutic Chemical code; DMT, Disease modifying therapy; ICD-10: 10th revision of the International Statistical Classification of Diseases and Related Health Problems; MS: multiple sclerosis, N/A: not applicable.

[1] The index date is defined as the date of initiation of Kesimpta or other DMT during study period.

Table 12-2 Exclusion criteria

Variable	ICD-10 code	Definition	Comments
Previous exposure to Kesimpta or other DMT before/at the index date[1]	-	-	See Table 12-3 for ATC
For Kesimpta cohort: prior Kesimpta exposure	-	-	See Table 12-3 for ATC
For the other DMT-initiator cohort: prior exposure to Kesimpta or to the same DMT	-	-	See Table 12-3 for ATC

Abbreviations: ATC, Anatomical Therapeutic Chemical code; DMT, Disease modifying therapy; ICD 10: 10th revision of the International Statistical Classification of Diseases and Related Health Problems;.

[1] The index date is defined as the date of initiation of Kesimpta or other DMT during study period.

Table 12-3 Anatomical Therapeutic Chemical (ATC) codes and risk windows of the study drugs

Variable[1,2]	ATC codes*	Risk window[3]	Exposure definition
Kesimpta (Ofatumumab)	L04AA52	180 days	See Section 7.3.1
Other DMT			For all other DMT, see Section 7.3.1
Alemtuzumab	L04AA34	120 days	
Azathioprine	L04AX01	25 hours	
Cladribine	L04AA40	180 days	
Cyclophosphamide	L01AA01	2.5 days	
Dimethyl fumarate (Tecfidera)	L04AX07	5 hours	
Diroximel fumarate (Vumerity)	L04AX09	5 hours	
Fingolimod	L04AA27	60 days	
Glatiramer acetate	L03AX13	Approx.1 day	
Interferon beta			
Avonex, Plegridy, Rebif, Extavia, Betaseron	L03AB02 L03AB07 L03AB08 L03AB13	Avonex 3 days Plegridy approx. 17 days Rebif approx. 22 days Extavia 1 day Betaseron 1 day	
Intravenous immunoglobulin	J06BA02	Approx. 180 days	
Leflunomide	L04AA13	75 days	
Methotrexate	L01BA01, L04AX03	Approx. 3.5 days	
Mitoxantrone	L01DB07	180 days	
Natalizumab	L04AA23	105 days	
Ocrelizumab	L04AA36	180 days	

Variable[1,2]	ATC codes*	Risk window[3]	Exposure definition
Ozanimod	L04AA38	Approx. 4.5 days	
Ponesimod	L04AA50	Approx. 7 days	
Siponimod	L04AA42	10 days	
Teriflunomide	L04AA31	95 days	

Abbreviations: ATC, Anatomical Therapeutic Chemical; DMT, Disease modifying therapy.

[1] The set of variables used for exposure ascertainment may differ, according to data availability. The final selection will be defined in the data source specific protocols, with the aim to achieve the most valid definition based on the available data.

[2] Additional DMTs could be added to the list if they become available in the study countries.

[3] The window of clearance of a multiple sclerosis disease modifying therapy is defined as five times drug half-life or the pharmacodynamic effect where applicable.

* Diagnosis and drug codes could be further adapted based on the availability and classification system in the local data sources and described in the localized protocols.

Table 12-4 Primary and secondary outcomes

Variable	ICD-10 Codes[1]	Definition	Distinction for analytical conventions
Primary objective			
Malignancies (overall) excluding non-melanoma skin cancer (NMSC)	C00-C97 except C44		Long-term outcome
Pre-defined malignancies		Records of a malignancy for each of the below	Long-term outcome
Skin neoplasms	C43-C44		
Melanoma of skin	C43		
Squamous cell carcinoma	C44.02, C44.12, C44.22, C44.32, C44.42, C44.52, C44.62, C44.72, C44.82, C44.92		
Basal-cell carcinoma	C44.01, C44.11, C44.21, C44.31, C44.41, C44.51, C44.61, C44.71, C44.81, C44.91		
In situ neoplasms	D00-D09		
Malignant neoplasm of other connective and soft tissue	C49		
Kaposi sarcoma	C46		
Lymphomas	C81-88, C96		

Variable	ICD-10 Codes[1]	Definition	Distinction for analytical conventions
Non-Hodgkin lymphoma	C82-86, C96		
B-cell lymphoma	C83.0, C83.3		
Burkitt's lymphoma	C83.7		
Hodgkin lymphoma	C81		
Mature T/NK-cell lymphomas	C84		
Mycosis fungoides	C84.0		
Multiple myeloma	C88, C90		
Leukemia	C91-95		
Breast	C50		
Brain, central nervous system	C70-72		
Colorectum	C18-21		
Lung	C33-34		
Cervix uteri	C53		
Corpus uteri	C54		
Ovary	C56		
Stomach	C16		
Liver	C22		
Pancreas	C25		
Kidney	C64-65		
Bladder	C67		
Esophagus	C15		
Thyroid	C73		
Prostate	C61		
Lip, oral cavity	C00-06		
Larynx	C32		
Nasopharynx	C11		
Gallbladder	C23		
Oropharynx	C09-10		
Hypopharynx	C12-13		
Testis	C62		
Salivary glands	C07-08		
Vulva	C51		
Penis	C60		

Variable	ICD-10 Codes[1]	Definition	Distinction for analytical conventions
Mesothelioma	C45		
Vagina	C52		
Late-onset infections		Records of an infection for each of the below	Long-term outcome
Overall late-onset infections			
Progressive multifocal leukoencephalopathy	A81.2		
HBV reactivation	No ICD codes. Use local codes if available		
Herpes zoster (Varicella zoster virus reactivation)	B02		
HIV	B20		
Tuberculosis	A15-A19		
Acute-onset infections		Records of an infection for each of the below	Short-term outcome
Overall acute-onset infections			
Nocardiosis	A43		
Listeriosis	A32		
Aspergillosis	B44		
Non-tuberculosis mycobacterial infections	A31		
Toxoplasmosis	B58		
Disseminated strongyloidiasis	B78.7		
Amoebiasis (lung and or brain)	A06.5, A06.6		
Leprosy	A30.1 - A30.5		
Cytomegalovirus infection	B25		
Coccidioidomycosis	B38		
Paracoccidioides infection	B41		
Histoplasmosis	B39		
Candidiasis	B37		
Cryptococcosis	B45		
Herpes simplex	B00		

Variable	ICD-10 Codes[1]	Definition	Distinction for analytical conventions
All fungal infections	B35-49		
Blastomycosis	B40		
Legionellosis	A48.1		
Salmonellosis	A02.1, A02.21 – A02.25		
Systemic Bartonellosis	B44.0		
Burkholderia infection	A24		
Isosporiasis	A07.3		
Parvovirus infection	B34.3		
Respiratory syncytial virus (RSV) infection	B97.4		
Varicella	B01.0 – B01.89		
Viral meningitis	A87		
Visceral leishmaniasis	B55.0		
Acute lower respiratory infections	J20-J22		
Urinary tract infection	N39.0		
Appendicitis	K35 – K37		
COVID-19 pneumonia	J12.81, J12.82		
Secondary objective			
Intestinal or bowel obstruction		Records of a diagnosis	Short-term outcome
Paralytic ileus	K56.0		
Intussusception	K56.1		
Volvulus	K56.2		
Intestinal obstructive adhesions	K56.5		
Other and unspecified intestinal obstruction	K56.6		
Ileus – unspecified	K56.7		
Ischemic stricture of intestine	K55.1		
Duodenal obstruction	K31.5		
Anal and rectal stenosis	K62.4		
Atony of colon	K59.8		
Diverticular disease of intestine with perforation or abscess	K57.8		

Variable	ICD-10 Codes[1]	Definition	Distinction for analytical conventions
Unspecified abdominal hernia with obstruction	K46		
Intussusception of appendix	K38.8		
Inflammatory bowel disease	K50-52		
Diverticular disease	K57		
Sarcoidosis	D86.0-D86.9	Records of a diagnosis	Long-term outcome
Suicidal ideation	R45.851	Records of a diagnosis	Short-term outcome
SAE	Any	Recorded as SAE	Short-term outcome

Abbreviations: HBV: hepatitis B virus; HIV: human immunodeficiency virus; ICD 10: 10th revision of the International Statistical Classification of Diseases and Related Health Problems.

[1] Diagnosis and drug codes could be further adapted based on the availability and classification system in the local data sources and described in the localized protocols.

Table 12-5 Censoring variables during follow-up

Variable	ICD-10 code[1]	Definition	Comments
Death	N/A	Date of death	-
End of data availability	N/A	Emigration, end of follow-up in the data, or analysis cut-off date	-

Abbreviations: ICD-10: 10th revision of the International Statistical Classification of Diseases and Related Health Problems; N/A: not applicable

[1] The codes could be further adapted based on the availability and classification system in the local data sources and described in the localized protocols.

Table 12-6 Other variables

Variable	Baseline characteristics	Candidate for PS	Definition[1]
Demographic characteristics			
Sex	Yes	Yes	-
Age at index date	Yes	Yes	-
Geographic area	Yes	Yes	-
Socio-economic status	Yes	Yes	-
Other variables			
Calendar year of index date[2]	Yes	Yes	-
Lifestyle characteristics			

Variable	Baseline characteristics	Candidate for PS	Definition[1]
Body mass index (height and weight)	Yes	Yes	-
Alcohol/substance abuse	Yes	Yes	Records of ICD-10 codes F10-F19, I42.6, G62.1-G62.2, K70, O35.4, O35.5
Smoking	Yes	Yes	Smoking status at baseline
MS disease characteristics			
MS disease duration at index date	Yes	Yes	Time since first MS diagnosis at the index date
Duration of MS treatment at the index date	Yes	Yes	-
MS type at index date (if available)	Yes	Yes	Classified as: Relapsing-Remitting MS (RRMS) Secondary Progressive MS (SPMS) Primary-Progressive MS (PPMS) Progressive-Relapsing MS (PRMS)
EDSS or proxy	Yes	Yes	Expanded Disability Status Scale (EDSS) score (including date of assessment) or a proxy measure if EDSS not available
Number of relapses (including date of onset), including glucocorticoid treatment (yes/no) within one year prior to enrolment and during follow-up	Yes	Yes	Definitions will be detailed in SAP and localized to each data source
John Cunningham virus antibody status positive and negative at index date	Yes	Yes	Definitions will be detailed in SAP and localized to each data source
Medical history, health care utilization			
Healthcare resource utilization	Yes	Yes	Definitions will be detailed in SAP and localized to each data source
Medical history, other comorbid conditions			Definitions will be detailed in SAP and

Variable	Baseline characteristics	Candidate for PS	Definition[1]
			localized to each data source
Cardiovascular disease	Yes	Yes	Records of ICD-10 diagnosis codes I00-I99
Respiratory disease	Yes	Yes	Records of ICD-10 diagnosis codes J00-J99
Gastrointestinal disease	Yes	Yes	Records of ICD-10 diagnosis codes K00-K93
Metabolic disease	Yes	Yes	Records of ICD-10 diagnosis codes E70-E90
Musculo-skeletal disease	Yes	Yes	Records of ICD-10 diagnosis codes M00-M99
Mental disorders	Yes	Yes	Records of ICD-10 diagnosis codes F00-F99
Autoimmune disease other than MS	Yes	Yes	Records of ICD-10 diagnosis codes
Vitamin B12 deficiency anemia due to intrinsic factor deficiency			D51.0
Drug-induced autoimmune hemolytic anemia			D59.0
Other autoimmune hemolytic anemias			D59.1
Other nonthrombocytopenic purpura			D69.2
Idiopathic thrombocytopenic purpura			D69.3
Sarcoidosis			D86
Thyrotoxicosis with diffuse goiter			E05.0
Autoimmune thyroiditis			E06.3
Type 1 diabetes mellitus			E10
Primary adrenocortical insufficiency			E27.1
Guillain–Barre syndrome			G61.0

Variable	Baseline characteristics	Candidate for PS	Definition[1]
Myasthenia gravis			G70.0
Iridocyclitis			H20
Raynaud's syndrome			I73.0
Inflammatory bowel disease			K50-K52
Primary biliary cirrhosis			K74.3
Autoimmune hepatitis			K75.4
Celiac disease			K90.0
Pemphigus			L10
Pemphigoid			L12
Psoriasis			L40
Alopecia areata			L63
Vitiligo			L80
Vasculitis limited to the skin, not elsewhere classified			L95
Rheumatoid arthritis, seropositive and other			M05-M06
Juvenile arthritis			M08-M09
Systemic connective tissue disorders			M30-M36
Ankylosing spondylitis			M45
Other specified disorders involving the immune mechanism, not elsewhere classified			D89.8
Immunodeficiencies	Yes	Yes	Record of ICD-10 diagnosis code D80-89
Prior infection	Yes	Yes	Records of late-onset and acute-onset infections as specified in Table 12-4
Previous malignancies[3]	Yes	Yes	Records of ICD-10 diagnosis codes C00-C97 and/or D00-D09
Family history of malignancy[3]	Yes	Yes	Records of ICD-10 diagnosis codes C00-C97 and/or D00-D09 in family history

Variable	Baseline characteristics	Candidate for PS	Definition[1]
SARS-COV-2 infection	Yes	Yes	Record of ICD-10 diagnosis code U07.1
Medication and exposure journey			Definitions will be detailed in SAP and localized to each data source
Kesimpta use	Yes	Yes	
Other DMT use (see Table 12-3)	Yes	Yes	
Concomitant medications	Yes	Yes	ATC code
Antiobesity preparations, excl. diet products			A08
Drugs used in diabetes			A10
Anabolic agents for systemic use			A14
Antithrombotic agents			B01, A01AD05, N02BA01
Antihypertensives			C02
Diuretics			C03
Beta blocking agents			C07
Calcium channel blockers			C08
Agents acting in the renin-angiotensin system			C09
Lipid modifying agents			C10
Intrauterine and intravaginal hormonal contraceptives			G02BB, G02BA03
Sex hormones and modulators of the genital system			G03
Corticosteroids			H02A, H02B, D07, S01BA, S01BB, S01CA, S01CB, S02B, S02C
Antibacterials and antiinfectives			J01, D06, S01A, S01C, S02A, S03C, G01
Antimycotics and antifungals			J02, D01
Antimycobacterials			J04
Antivirals			J05, D06BB, S01AD
Immune sera and immunoglobulins			J06

Variable	Baseline characteristics	Candidate for PS	Definition[1]
Antineoplastic agents			L01
Endocrine therapy			L02
Immunostimulants			L03
Immunosuppressive agents			L04
Antiepileptics			N03
Hypnotics and sedatives, Anxiolytics			N05C, N05B
Antidepressants			N06A, N06CA
Antiprotozoals			P01
Anthelmintics			P02
Vaccines (name and date)	Yes	Yes	ATC codes J07
Radiation exposure	Yes	Yes	Definitions will be detailed in SAP and localized to each data source

Abbreviations: ATC: Anatomical Therapeutic Chemical; DMT: multiple sclerosis disease modifying therapy; ICD 10: 10th revision of the International Statistical Classification of Diseases and Related Health Problems; MS: multiple sclerosis; PS: propensity score; SAP: Statistical Analysis Plan.

[1] Diagnosis, drug codes and other definitions could be further adapted based on the availability and classification system in the local data sources and described in the localized protocols.

[2] not directly related to outcomes however can be a proxy to measure change in health care standard

[3] Partially collected in MS quality registers

12.4 Annex 4 – Database selection – Feasibility assessment findings

A feasibility assessment was performed, Q4-2021, to identify the database suitable to address the question of interest. The main criteria included but were not limited to:

a. capability to identify MS patients,
b. size of the active MS population,
c. length of the average follow-up,
d. ability to assess MS type,
e. availability of overall and specific malignancies and infections,
f. availability of DMT exposures.

In addition to the data sources selected as described above (Table 7-2), Table 12-7 provides the data sources not selected along with the rational for the decision.

Table 12-7 List of data sources not selected with rationale

Study population and country	Data source name	Rationale for exclusion
US	IQVIA PharMetrics Plus IBM MarketScan OptumInsight	Limited amount of follow-up time, not allowing to assess long-term outcomes
Canada	ICES Administrative Data Holdings for Ontario Alberta Health Data Asset Directory Manitoba Centre for Health Policy (MCHP)	Drugs dispensed in hospital not captured i.e., high risk for exposure misclassification
Denmark	Danish Multiple Sclerosis Treatment Register (SCBH)	SCBH records only RRMS
Finland	Finnish Population Information System ePrescription register Register for Reimbursed Medications Care Register for Health Care (HILMO; secondary care) Register of Primary Health Care Visits (AvoHILMO) National Infectious Diseases Register Finnish Cancer Registry Cause of Death Register Finnish MS register (StellarQ Ltd)	Relatively small active MS population
Norway	National Population Register Norwegian Prescription Database (NorPD) Norwegian Patient Registry (NPR-Nor) Cancer Registry of Norway Cause of Death Registry Norwegian Immunisation Registry (SYSVAK) Norwegian MS Registry and Biobank	Kesimpta not re-imbursed

Study population and country	Data source name	Rationale for exclusion
Germany	NeuroTransData (NTD) Multiple sclerosis management system 3D (MSDS 3D)	Not suitable for secondary use of data

12.4.1 Detailed description of the selected data sources

12.4.1.1 Denmark

12.4.1.1.1 Denmark Civil registration system (CPR)

The Danish Civil Registration System or Central Person Register (CPR) copy is operated by the Danish Ministry of the Interior. The Danish Civil Registration System was introduced in 1968, information on demographics (age, sex, geographical region), migration and vital statistics data (date of birth and date of death) has been registered electronically daily for all Danish residents. Every individual in Denmark is provided with a unique personal identification number (CPR number) at birth or upon emigration which allows for follow-up until death or emigration. The CPR number forms the basis for the precise, deterministic linkage of individual-level data between all patient-level registers and databases in Denmark, allowing the creation of a study database with individual-level data for any given study. The CPR number also allows for family linkage of data.

12.4.1.1.2 Danish Register of Medicinal Product Statistics (RMPS)

The Danish RMPS contains patient-level data on all prescription drugs filled by patients at community pharmacies. The register contains information on the date of purchase, item number, product name, ATC code, strength per unit, quantity of the World Health Organization's defined daily doses per package and number of packages filled. Data are available from 1995 and onwards. Since April 2004, information on medical indication for prescription and daily prescribed dose by physician has also been available, but completeness and validity are affected by a non-compulsory obligation to record this information. At Danish Health Data Authority, the register is updated once a month, typically with a lag of approximately two months. If the project requires that data are stored at Statistics Denmark (DST) because of link to socio-economic registers, the Danish RMPS is only updated twice a year after end of June and December with a lag of approximately 5-6 months.

12.4.1.1.3 Danish National Hospital Medication Register (SMR)

The SMR contains information on drugs administered to patients while admitted to hospital or during outpatient visits: the date and time of administering the drug, the dose administered via number of units and strength per unit, product name, ATC code, and department information. Thus, this register complements the Danish RMPS. Data is captured since May 2018 and the register is updated monthly with a lag of approximately two months.

12.4.1.1.4 Danish National Patient Register (NPR-Den)

The Danish National Patient Register (NPR-Den) was established in 1977 and is considered to have high completeness and validity. Since 2007 the register has included information on all

patients in Danish hospitals including private hospitals, however, reporting from private hospitals and clinics are not considered complete. The validity and coverage are very high, but both the content and the definitions of single variables have changed over time (Schmidt et al, 2015).

The register uses Health Care Classification System (Danish, Sundhedsvæsenets Klassifikations System) for different classifications. The NPR-Den includes the following information: CPR number, local municipality, admission and discharge information, the date of any incidents over the course of an illness, diagnosis (ICD-10), examinations and treatment information - including surgery coded with Danish version of the Nordic Medico-Statistical Committee Classification of Surgical Procedures codes as well as supplementary information regarding births. The register is updated once a month, typically with a lag of approximately two months. However, when data from socio-economic registers must be included, too, the complete linked datasets must be stored at DST and, at DST the NPR-Den is updated once a year after end of December with a lag of approximately ten months.

12.4.1.1.5 Registry of Laboratory Results for Research (RLRR)

Information on laboratory tests at the country's large clinical biochemical and clinical immunological laboratories and the result of the tests are collected in a National Laboratory Database on a daily basis. Tests performed in general practice and sent to a hospital laboratory for analysis are also included. This information is transferred to the RLRR. If a patient has, proactively, denied consent to exchange lab results, the results are not transferred to the RLRR.

Two types of tests are included: a) analysis of blood, urine, joint fluids, and spinal fluids for the purpose of preventing, diagnosing, and controlling the treatment of human diseases; b) blood type determination and examination of blood by e.g., pregnancy, immune disorders, and certain infections. The register contains information on type of test, value, unit, and date of sampling.

Data from the five regions in Denmark has gradually been included in the register since late 2013. In addition, some historical data from before 2013 has also been transferred to the register. Completeness is not reported but expected to be very high. Full national coverage from alle regions was accomplished in 2020.

Only laboratory tests encoded with Nomenclature for Property and Unit (NPU) terminology (that is codes beginning with either NPU or Danish version of it called DNK) is transferred from National Laboratory Database to the RLRR (currently approximately 95% of all laboratory tests).

12.4.1.1.6 Danish Cancer Register

The Danish Cancer Register contains patient-level data on incident cancers in the Danish population. Data must be reported to the register according to an executive order by The National Board of Health. Data are available from 1943 and onwards, however, reporting was voluntary until March 1987. The register contains information on date of diagnosis, location, classification of malignant tumors (TNM: a classification system to describe the amount and spread of malignant tumors in a patient’s body) classification, and histological classification of cancer. Information from the NPR-Den, the Cause of Death Register and the Danish Pathology Register is used to increase completion and accuracy of data collection in the process of

updating the Cancer Registry. Validation of the Cancer Register and selected clinical cancer databases was reported in 2012 by the authorities with histological verification of 92-96% of the tumors, and even close to 100% in some cancer types, including melanoma, breast, lung, and colon cancer. The register is updated once a year after end of December with a lag of approximately 13 months. Since 2004, ICD-O-3 has been used as the classification system. The ICD-7 classification was used until 1977. Between 1978 and 2003, ICD-O-1 was used, but has since then been converted to ICD-O-3.

12.4.1.1.7 Danish Cause of Death Register

The Danish Cause of Death Register contains information on date and cause of death. When a person dies in Denmark, a death certificate is filled by a medical doctor. The information from the death certificate, including place of death, information about any autopsy, and municipality of residence is transferred to the Cause of Death register. Date of death is available from 1970 and onwards. Causes of death have been recorded since 1970, and since 2002 have been post-processed with the Automated Classification of Medical Entities system to ensure a uniform recording of the underlying cause of death over time. The register is updated once a year after end of December with a lag of approximately 13 months. The ICD-8 was used up until 1994 and since 1994, ICD-10 has been used as classification system.

If needed, it is possible to get an unvalidated version of the Cause of Death Register, while waiting for the validated version of the death in the most recent calendar year. These unvalidated causes of death can later be replaced with the validated version.

12.4.1.1.8 The National Health Insurance Service Register (NHISR)

The register contains information on services supported by public health insurance and provided by general practitioners and specialists in private practice outside the secondary setting at hospitals (Andrade et al, 2021). Specialists in private practice include anesthesiologists, psychologists, psychiatrists, pediatricians, gynecologists, ophthalmologists, rheumatologists, radiologists, chiropodists, chiropractors, neurologists, orthopedic surgeons, plastic surgeons, physiotherapists, occupational therapists, and dentists. Data is available from 1990 and onwards and the register is updated once a month, typically with a lag of approximately 2 to 4 months.

12.4.1.1.9 Danish Covid-19 Surveillance Database

Database is maintained by the Statens Serum Institute and includes information of both positive and negative COVID-19 tests as well as COVID-19 vaccinations (all COVID-19 vaccinations, and type of vaccines as recoded in July 2021). Data can be applied through Danish Health Data Authority.

12.4.1.1.10 Danish Multiple Sclerosis Registry (DMSR)

Danish Multiple Sclerosis Registry (DMSR) was established in 1956 and contains data on all Danish residents diagnosed with MS from 1948 to present. Treatment history for patients has been registered from 1996. AEs from treatments have been reported from the beginning and have been optimized to contain more details since 2018. DMSR contains information on demographics, clinical and paraclinical data, EDSS, Magnetic Resonance Imaging (MRI), relapses, and AEs. In the beginning of 2020, the database included data from 30,023 patients of

whom 16,515 were alive with CIS and MS. The mean annual number of new cases receiving an MS diagnosis was 649 per year in the period 2010 to 2019. Nearly 50% of the patients included in the registry are receiving Dimenthyltryptamine. A 2012 validation study emphasized the completeness and reliability of the DMSR (Mason et al, 2012).

12.4.1.2 Sweden

12.4.1.2.1 Total Population Register

This register is managed by the Statistics Sweden and contains information on all persons registered in Sweden. The register was established in 1968 and contains data on life events including birth, death, name change, marital status, family relationships and migrations, residence, citizenship and country of birth. The information in the register is updated daily. All citizens in Sweden are assigned a personal identification number which is used in all of the national registries for identification and can be used to link information across registries.

12.4.1.2.2 Swedish Prescribed Drug Register (SPDR)

The Prescribed Drug Register was established in July 2005 and is maintained by the National Board of Health and Welfare. It contains all medicines that are collected against prescriptions at pharmacies, but also information on collected preferential consumables. Furthermore, it includes information on the patient (sex, age, place of registration) and product information (including ATC code, drug name, strength, pack size). Monthly data are available with a lag time of 1-2 months. Annual data (data up until December of previous year) is available in April the following year. Validity of the data is considered to be high.

12.4.1.2.3 National Patient Register (NPR-Swe)

The NPR-Swe was established in 1964, with complete coverage starting from 1987. The registry comprises data on healthcare encounters in in-patient (hospital) and outpatient specialist care given in public hospitals. The coverage for in-patient care is complete since 1987 (ICD-10 from 1998) and specialized outpatient care from 2001 onwards. Outpatient care data is included from a proportion of private hospitals, but the data is partly missing, while for public hospitals it is close to 100%. The NPR contains information on, e.g., details of hospitalization, all disease areas, hospital and hospital department, visit date, diagnosis, diagnosis date, comorbidities at diagnosis, (ICD 9 or -10 codes), procedures, and patient demographics.

Once a month, each of the 21 county councils in Sweden and private caregivers deliver information to the register. Quarterly updates are available, with some missingness to be expected. Full updates are available once a year with data up to end of December of previous year. The data are available for research with ~6-9 months delay.

Information included in the registry:

- All completed care sessions in in-patient care since 1964 (comprehensive from 1987);
- Data on patients treated by physicians in specialized outpatient care since 2001; and
- Information on emergency waiting times and emergency operations since 2016.

12.4.1.2.4 Swedish Cancer Register

Swedish Cancer Register was established in 1958 and is managed by the National Board of Health and Welfare. The registry records every newly discovered primary tumor and all blood cancer cases diagnosed by clinical, morphological, or other laboratory examinations as well as cases diagnosed at autopsy. Since the early 1980's, there are six regional registries associated with the oncological centers in each medical region of Sweden where the registration, coding and major check-up and correction work is performed. The register includes information on patient characteristics, tumor characteristics, diagnosis, and date and cause of death. Data is updated on an annual basis with a lag time of approximately 12 months. Applications for data access is sent to the National Board of Health and Welfare and processed by a pre-defined data permit process, preceded by ethical review by the Swedish Ethical Review Authority.

12.4.1.2.5 Swedish Cause of Death Register

The Swedish Cause of Death Register is a high-quality virtually complete register of all deaths in Sweden since 1952. The main strengths of the Swedish Cause of Death Register are the high completeness and long history, with data electronically available since 1952. The quality of the whole register has not been checked since 1995. The data are updated yearly and includes 4-6 months lag for date of death. Cause of death data for previous year is available for research in April next year.

12.4.1.2.6 Swedish Infectious Disease Register (SmiNet)

SmiNet is a system for communicable disease surveillance in Sweden, established in 2004. Practicing doctors and laboratories can report individual cases of notifiable communicable diseases in Sweden on SmiNet. Suspected cases of COVID-19 are reported on SmiNet.

12.4.1.2.7 Swedish National Quality Registry for Neurological Care

The Swedish National Quality Registry for Neurological Care includes eight disease groups within neurology, one being MS. The registry includes approximately 80% percent of all MS patients in the country. In addition to demographic data and medical history (e.g., age of onset, initial symptoms, and disease course type) outcome of diagnostic investigation with MRI and cerebrospinal fluid exams, measurement of disease activity, rating scales of disability, quality of life, fatigue and cognition and information on treatments are available in the registry.

12.4.1.3 France

12.4.1.3.1 French Observatory of Multiple Sclerosis (OFSEP)

The French Observatory of Multiple Sclerosis (Observatoire Français de la Sclérose en Plaques) (OFSEP) was established in 2003. It is a national register actively covering data from 37 centers and representing about a quarter to a third of all MS patients in France. The centers use an e-record software, the European Database for Multiple Sclerosis (EDMUS), to record health information. The OFSEP systematically collects information on personal, sociodemographic, clinical, paraclinical, and therapeutic data twice a year from EDMUS, in respect of patient confidentiality rules.

The OFSEP database is frequently used for research purposes and information can be accessed through collaboration with a research team of the OFSEP. The OFSEP will perform data analysis locally.

12.4.1.4 Italy

12.4.1.4.1 Italian MS Register (IMSR)

The IMSR was established in 2014, and contains personal, clinical and healthcare data of patients with MS. The IMSR collects data from 140 out of 236 contacted centers who have declared the willingness to participate (last update May 2018), and 103 completed their ethics committee process for approval and are ready to participate in the data collection. The IMSR includes data on approximately 60% of MS population in Italy representing all geographic regions in the country. On May 2018, 72 MS centers effectively contributed uploading data to the registry and the contribution level has increased during the time. There are specific regions, e.g. Apulia, Emilia-Romagna, Liguria, Tuscany, where patient-level data can be linked to administrative database.

Healthcare data is regularly used for research purposes and can be accessed through the IMSR management committee and specific agreement should be signed between [redacted] IMSR and University of Bari Alda Moro. The data extraction depends on report frequency and data management procedures may take up to months based on analysis complexity. The IMSR will perform data analysis locally.

12.4.1.5 MSBase

MSBase is a multi-center, multi-country longitudinal observational MS database and was established and dedicated for sharing, tracking and evaluating outcomes data in MS and other central nervous system demyelinating diseases. The register comprises information on patient characteristics, including family and patient's medical history; clinical evaluation of patients (e.g., MS diagnosis, symptoms, information on safety and other events, diagnostic and laboratory tests); and MS specific and symptomatic and non-pharmacological treatment.

In MSBase, the patient enrolment started in 2003, and by December 2021, MSBase has accumulated over 78,865 patient records from 39 participating countries. MSBase has developed minimum essential and desirable fields for a dataset that is captured with a specific data entry software. Pseudonymized data from data entry software in MSBase is uploaded real-time or monthly to the central database, based on-site preference. The annual lost to follow-up rate is 4%.

The MSBase registry requires each stakeholder to maintain strict confidentiality over individual patient, caregiver, physician, and clinic data. Individual neurologists have access only to their own patient data and pooled results of aggregate outcomes in the registry. The MSBase will perform data analysis locally and the Sponsor will have access to aggregate reports only, not to individual patient, physician, or hospital data.