TITLE:	**MULTISOURCE SURVEILLANCE STUDY OF PREGNANCY AND INFANT OUTCOMES IN OCRELIZUMAB-EXPOSED WOMEN WITH MULTIPLE SCLEROSIS**
PROTOCOL NUMBER:	BA39732
VERSION NUMBER:	1.0
AUTHOR:	[redacted] F. Hoffmann-La Roche Ltd Grenzacherstr. 124 4070 Basel, Switzerland
DATE FINAL:	See electronic date stamp below

FINAL PROTOCOL APPROVAL

Date and Time (UTC)	Title	Approver's Name
27-Jan-2020 19:04:28	Deputy EU QPPV	[redacted]
27-Jan-2020 20:03:10	CMD	[redacted]
27-Jan-2020 21:29:17	Company Signatory	[redacted]

CONFIDENTIAL

EU PAS REGISTER NUMBER:	Study not registered
ACTIVE SUBSTANCE:	L04AA36 (ocrelizumab)
STUDIED MEDICINAL PRODUCT:	OCREVUS®
PRODUCT REFERENCE NUMBER(S):	RO4964913
PROCEDURE NUMBER(S):	EMEA/H/C/004043; IND 100,593 BLA 761053
JOINT PASS:	No
RESEARCH QUESTION AND OBJECTIVES:	To assess and characterize pregnancy and infant outcomes of women with multiple sclerosis exposed to ocrelizumab during the 6 months before the estimated date of conception or at any time during pregnancy
COUNTRIES OF STUDY POPULATION:	United States, Denmark
MARKETING AUTHORIZATION HOLDER (MAH):	Roche Registration Ltd 6 Falcon Way Shire Park Welwyn Garden City AL7 1TW United Kingdom
MAH CONTACT PERSON:	[redacted] F.Hoffmann-La Roche AG Hochstrasse 16 CH-4053 Basel

1. TABLE OF CONTENTS

1. TABLE OF CONTENTS 3
2. LIST OF ABBREVIATIONS 6
3. RESPONSIBLE PARTIES 8
4. ABSTRACT/SYNOPSIS 11
5. AMENDMENTS AND UPDATES 17
6. MILESTONES 17
7. RATIONALE AND BACKGROUND 20
   - 7.1 Study Background 21
8. RESEARCH QUESTION AND OBJECTIVES 23
9. RESEARCH METHODS 23
   - 9.1 Study Design 25
   - 9.2 Setting 25
   - 9.2.1 Eligibility Criteria 26
   - 9.2.2 Study Period 26
   - 9.2.3 Follow-Up 27
   - 9.3 Variables 27
   - 9.3.1 Exposure 27
   - 9.3.2 Outcomes 29
   - 9.3.3 Covariates 32
   - 9.4 Data Sources 39
   - 9.4.1 HealthCore Integrated Research Database 40
   - 9.4.2 Optum Dynamic Assessment of Pregnancies and Infants (DAPI) 41
   - 9.4.3 IBM MarketScan Commercial Claims and Encounters Database 41
   - 9.4.4 Danish National Health Databases and the Multiple Sclerosis Registry 42
   - 9.5 Study Size 44
   - 9.6 Data Management 47
   - 9.7 Data Analysis 48

9.7.1 Study Cohorts 48
9.7.2 Descriptive Analyses 52
9.7.3 Measures of Frequency 52
9.7.4 Measures of Association 52
9.7.5 Missing Data 53
9.7.6 Statistical Analyses 53
9.7.7 Data Integration 58
9.7.8 Subgroup Analyses 58
9.7.9 Sensitivity Analyses 58
9.8 Quality Control 59
9.9 Limitations of the Research Methods 59

10. PROTECTION OF HUMAN SUBJECTS 61
10.1 RTI International 62
10.2 HealthCore Integrated Research Database 62
10.3 Optum Dynamic Assessment of Pregnancies and Infants 62
10.4 IBM MarketScan Commercial Claims and Encounters Database 63
10.5 Danish National Health Databases and Danish Multiple Sclerosis Registry 63
10.6 Other Good Research Practice 63

11. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS 63

12. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS 64

13. REFERENCES 65

LIST OF APPENDICES

Appendix 1 Contact Details for Collaborating Institutions 73
Appendix 2 ENCePP Checklist for Study Protocols 74
Appendix 3 Algorithms to Identify Key Study Endpoints 82
Appendix 4 Characteristics of Study Data Sources 85

LIST OF TABLES

Table 1 Study Milestones 19
Table 2 Ocrelizumab and Other Multiple Sclerosis Disease-Modifying Therapies 28
Table 3 Covariates 33
Table 4 Published Algorithms to Identify Relapses of Multiple Sclerosis in Claims Data 38
Table 5 Study Sizes Needed to Have Probability of 0.8 That the Upper Limit of the 95% Confidence Intervals Will Be Below Selected Thresholds 45
Table 6 Estimated Number of Pregnancies Exposed to Ocrelizumab in Women With Multiple Sclerosis in Selected US Data Sources, per Year 46
Table 7 Study Cohorts 49
Table 8 Published Algorithms to Identify Cases of Multiple Sclerosis in Claims Data 50
Table 9 Proposed Statistical Analyses 55

LIST OF FIGURES

Figure 1 Study Timeline 17

2. LIST OF ABBREVIATIONS

Abbreviation	Definition
ATC	Anatomical therapeutic chemical
CDC	Centers for Disease Control and Prevention
CFR	Code of Federal Regulations
CI	Confidence interval
CNS	Central nervous system
CSR	Clinical study report
DAPI	Dynamic Assessment of Pregnancies and Infants
DDD	Defined daily dose
DMSR	Danish Multiple Sclerosis Registry
DMT	Disease-modifying therapy
DVP	Data validation plan
EDSS	Expanded Disability Status Scale
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU PAS Register	European Union electronic register of post-authorization studies
EUROCAT	European Surveillance of Congenital Anomalies program
FDA	Food and Drug Administration
FSS	Functional Systems Score
GPP	Guidelines for Good Pharmacoepidemiology Practices
GVP	Guideline on Good Pharmacovigilance Practices
HIPAA	Health Insurance Portability and Accountability Act
HIRD	HealthCore Integrated Research Database[SM]
ICD-9	*International Classification of Diseases, 9th Revision*
ICD-9-CM	*International Classification of Diseases, 9th Revision, Clinical Modification*
ICD-10	*International Statistical Classification of Diseases and Related Health Problems, 10th Revision*
ICD-10-CA	*International Classification of Diseases, 10th Revision, Canadian Modification*
ICD-10-CM	*International Classification of Diseases, 10th Revision, Clinical Modification*
ICH	International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use
IgG1	Immunoglobulin G1

Abbreviation	Definition
IM	Intramuscular
IRB	Institutional review board
ISPE	International Society for Pharmacoepidemiology
IV	Intravenous
MarketScan	IBM MarketScan Commercial Claims and Encounters Database
MCM	Major congenital malformation
MRI	Magnetic resonance imaging
MS	Multiple sclerosis
MSIF	Multiple Sclerosis International Federation
NA	Not applicable
NDI	National Death Index
NMSS	National Multiple Sclerosis Society
NPV	Negative predictive value
PADER	Periodic adverse drug experience report
PAS	Post-authorization study
PASS	Post-authorization safety study
PBRER	Periodic benefit-risk evaluation report
PHI	Protected health information
PMR	Post-marketing requirement
PPMS	Primary progressive multiple sclerosis
PPV	Positive predictive value
Qn	Quarter of the year
RMS	Relapsing forms of multiple sclerosis
RRMS	Relapsing-remitting multiple sclerosis
RTI-HS	RTI Health Solutions
SAP	Statistical analysis plan
SC	Subcutaneous
SGA	Small for gestational age
STORK	Systematic Tracking of Real Kids
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TORCH	Toxoplasmosis, Other (syphilis, varicella-zoster, parvovirus B19), Rubella, Cytomegalovirus (CMV), and Herpes
US	United States
WHO	World Health Organization

3. RESPONSIBLE PARTIES

Each research partner under contract with Roche, as detailed further below, will share responsibility for the study design and conduct of the study and carry out all aspects of study implementation, including but not limited to:

- Obtaining local institutional review board and ethics approvals, as required;
- Interacting with other registers to achieve linkage with required data sets;
- Contributing to the development of the ocrelizumab uptake monitoring plan, statistical analysis plan (SAP), and study data validation plan (DVP);
- Managing raw data files;
- Obtaining ocrelizumab uptake monitoring data; and
- Preparing the analytical data sets following the protocol, the ocrelizumab uptake monitoring plan, the SAP, and any activities relating to validation of the study endpoints.

In addition to the responsibilities listed above for research partners, RTI Health Solutions serves as the coordinating center, with responsibilities including:

- Facilitating communication among all research partners, Roche, and the independent scientific advisors;
- Drafting and coordinating the development of common documents (e.g., the protocol, SAP, and reports); and
- Conducting the meta-analysis of aggregated results from individual data sources.

The investigators at RTI Health Solutions are responsible for conducting analyses in the IBM MarketScan Commercial Claims and Encounters Database. The investigators at HealthCore are responsible for conducting analyses in the HealthCore Integrated Research Database℠; investigators at Optum are responsible for conducting analyses in the Optum Dynamic Assessment of Pregnancies and Infants™ (DAPI); and investigators at [redacted] in Denmark are responsible for conducting analyses in the Danish National Health Databases and the Danish Multiple Sclerosis Registry.

Independent scientific advisors, listed below, will provide scientific and technical advice and recommendations throughout the project.

Protocol Development Responsible

F. Hoffmann-La Roche Ltd 4070 Basel, Switzerland
[redacted] Neuroscience Pharma Development Personalised Health Care - Data Science [redacted] DS Global Safety Science CNS

RTI Health Solutions—Barcelona Diagonal 605, 9-1 08028 Barcelona, Spain Phone: [redacted]
[redacted]

Coordinating Center

RTI Health Solutions—Barcelona Diagonal 605, 9-1 08028 Barcelona, Spain Phone: [redacted]

External Experts

Consultant
[redacted] MA Address: [redacted] MA 02446, USA

Consultant
[redacted] Denmark

4. ABSTRACT/SYNOPSIS

TITLE:	**MULTISOURCE SURVEILLANCE STUDY OF PREGNANCY AND INFANT OUTCOMES IN OCRELIZUMAB-EXPOSED WOMEN WITH MULTIPLE SCLEROSIS**
PROTOCOL NUMBER:	BA39732
VERSION NUMBER:	1.0
DATE OF PROTOCOL:	See electronic date stamp on the cover page

Rationale and Background

Multiple sclerosis (MS) is a chronic inflammatory demyelinating disease of the central nervous system, thought to be a complex interaction of genetic susceptibility and environmental factors. Its prevalence is highest in North America and Europe. Multiple sclerosis is more frequent in women than men, with an overall sex ratio of 2 to 3 women for every 1 man, except in patients with primary progressive MS, where the frequency is similar in both sexes. Multiple sclerosis is commonly diagnosed during reproductive years. About 85% of people with MS have relapsing-remitting MS, a form of MS with exacerbations or relapses during which new symptoms appear or previous symptoms worsen and periods of partial or complete remission. If left untreated, most patients who are diagnosed with relapsing-remitting MS will eventually progress to a chronic form characterized by progressive worsening of neurologic function over time with occasional relapses, which is categorized as secondary progressive MS. Approximately 15% of people with MS are diagnosed with primary progressive MS (PPMS), a form of MS with steadily worsening neurologic symptoms from the onset of disease and no clear relapses.

OCREVUS® (ocrelizumab) was approved by the United States (US) Food and Drug Administration (FDA) on March 28, 2017, for the treatment of adult patients with relapsing forms of MS (RMS) and PPMS. Subsequently, OCREVUS® was approved in the EU, Switzerland, Australia, Canada, and other countries.

Ocrelizumab is a recombinant humanized monoclonal immunoglobulin G1 (IgG1) antibody that selectively targets CD20-expressing (CD20+) B cells. Two identical, randomized, active-controlled studies (OPERA I [Study WA21092] and OPERA II [Study WA21093]) have demonstrated superior efficacy outcomes versus interferon beta-1a in RMS; one randomized, placebo-controlled study (ORATORIO [Study WA25046]) demonstrated superior efficacy in PPMS versus placebo. Results of these studies show that depletion of CD20+ B cells leads to a significant impact on a broad range of clinical measures of disease, including disability progression, in addition to an impact on magnetic resonance imaging (MRI) outcomes related to disease progression and reflective of neural tissue loss, thus further supporting the hypothesis that B cells are

central to the pathogenesis of both relapsing and primary progressive MS. Ocrelizumab has demonstrated a favorable safety profile in patients with RMS or PPMS. The proportion of patients with adverse events (AEs) was similar in patients treated with ocrelizumab and patients treated with interferon beta-1a (both 83.3%) or placebo (95.1% vs. 90.0%). The proportion of patients experiencing a serious adverse event was similar between ocrelizumab and the comparator groups—in RMS: 6.9%, ocrelizumab, and 8.7%, interferon beta-1a; in PPMS: 20.4%, ocrelizumab, and 22.2%, placebo.

Ocrelizumab is a humanized monoclonal antibody of an immunoglobulin G1 subtype, and immunoglobulins are known to cross the placental barrier. B-cell levels in human neonates following maternal exposure to ocrelizumab have not been studied in clinical studies. There are no adequate and well-controlled data from studies in pregnant women; however, transient, peripheral B-cell depletion and lymphocytopenia have been reported in infants born to mothers exposed to other anti-CD20 antibodies during pregnancy. It is not known whether ocrelizumab can affect pregnancy outcomes or infant outcomes in humans.

Roche proposes a non-interventional multidatabase post-marketing safety study to assess pregnancy-related safety data from women with MS exposed to ocrelizumab. This study will be conducted to fulfill part of the FDA post-marketing requirements (PMR 3194-4) for approval of ocrelizumab in the US. The study is listed as an additional pharmacovigilance activity to address the missing information on drug use during pregnancy and lactation in the proposed European Union Risk Management Plan.

Research Question and Objectives

To assess and characterize pregnancy and infant outcomes of women with MS exposed to ocrelizumab during the 6 months before the estimated date of conception or at any time during pregnancy.

Specifically, the objectives are as follows:

- To estimate the frequency of selected adverse pregnancy outcomes in women with MS exposed to ocrelizumab during the defined exposure window; i.e., spontaneous abortions, stillbirths, elective abortions, preterm births, C-sections, and urinary and other infections in pregnancy
- To estimate the frequency of selected adverse fetal/neonatal/infant outcomes at birth and up to the first year of life in infants from pregnancies in women with MS exposed to ocrelizumab; i.e., major congenital malformations, small for gestational age, adverse effects on immune system development (e.g., severe infectious disease in the first year of life)

  To compare the frequency of each safety event of interest between ocrelizumab-exposed pregnant women with MS and two comparison cohorts: (1) primary comparison cohort: pregnancies in women with MS who have not been exposed to ocrelizumab (overall and in two strata—pregnancies exposed to any approved non-ocrelizumab disease-modifying therapies [DMTs] for the treatment of MS or that

may be approved during the study period [subcohort 1a] and pregnancies not exposed to approved DMTs [subcohort 1b]) and (2) secondary comparison cohort: pregnancies in women without MS who have not been exposed to ocrelizumab

Study Design

This will be an observational cohort study of ocrelizumab-exposed pregnancies and two matched comparison cohorts using multiple sources of prospectively collected secondary data. The study will be conducted in existing population-based health care databases and registries. The proposed data sources include health care claims from the US and country-level registries from Denmark. Analyses will follow a common protocol and statistical analysis plan; results will be integrated via meta-analysis. Drug uptake will be monitored annually to determine when the main analysis should start, i.e., when a sufficient number of exposed pregnancies has accumulated (see Study Size). The earliest conception dates in the study data are expected to be in 2017. The study period will start with the first ocrelizumab prescription in each database (expected in 2017 in the US) and will end at the latest date for which data are available for the study from each data source (expected in Q2 2028).

The following cohorts will be assembled: ocrelizumab-exposed pregnancies in women with MS (exposed cohort), pregnancies not exposed to ocrelizumab in women with MS, and pregnancies not exposed to ocrelizumab in women without MS.

Population

Eligibility: The study population will include women from the three study cohorts and their children born during the study period. In each data source, study eligibility for mothers will require continuous enrollment with both medical and pharmacy benefits in the 6 months before the estimated beginning of pregnancy and throughout pregnancy. For children, eligibility criteria are inclusion of their mother in one of the three study cohorts (i.e., successful mother-child linkage is required) and continuous enrollment covering outpatient care and hospitalizations during follow-up (disenrollment will trigger end of follow-up).

For pregnancy outcomes, linkage to an infant is not required (i.e., pregnancies not linked to infants will be retained). For infant outcomes, linkage between the mother and infant is required.

Follow-up: Follow-up of women will start at the estimated beginning of pregnancy and will finish at the end of pregnancy; follow-up of infants will start at birth and finish at 1 year of age. For mothers and infants, follow-up will finish at the earliest of death, disenrollment from the data source, or end of the study period. For each outcome of interest that can occur multiple times, follow-up for that outcome will stop at its first occurrence (e.g., urinary tract infections in pregnancy, infections requiring hospitalization in pregnancy).

Variables

Information may vary in the level of detail across data sources.

Exposure

The exposure of interest is ocrelizumab. The exposure window will be from 6 months before conception (to take into account the product's half-life and the US prescribing information at the time of approval, March 2017) to the end of pregnancy; to the end of the first trimester for analyses on congenital malformations; to the end of the first trimester or end of pregnancy if prior to end of first trimester, for analyses on spontaneous abortions and elective terminations; or to the date of outcome(s) that may occur before the end of pregnancy (e.g., infections during pregnancy).

The sources for ascertainment of medication use will be pharmacy dispensing claims in claim databases and treatment information recorded in the MS Registry (possibly supplemented with data on dispensed prescriptions) in Denmark. The timing of medication use relative to the beginning of pregnancy will be determined or estimated from available data in the data sources.

Outcomes

- Spontaneous abortion (pregnancy loss at < 20 completed weeks)
- Fetal death/stillbirth (≥ 20 completed weeks)
- Elective termination: the reason for termination (e.g., therapeutic abortion, abnormal findings in fetus in prenatal tests, ectopic pregnancy) will be ascertained to the extent to which data are available
- Preterm delivery (live birth at < 37 completed weeks)
- C-section
- Urinary tract infection in pregnancy
- Infections requiring hospitalization during pregnancy
- Major congenital malformations in the infants
- Minor congenital malformations in the infants to the extent available, including a category for unspecified cardiac defects
- Small for gestational age
- Adverse effects on the immune system of the infants in the first year, comprising hospitalizations due to infectious diseases, cancer, and vaccine-preventable diseases and vaccine-associated poliomyelitis
- Infant growth and development to the extent available

Covariates

These variables were selected to provide a general description of the health and characteristics of the study cohorts and to explore potential confounding.

- Maternal characteristics at conception: age, education, obesity, medical history (including MS subtype [in Denmark] and other available information), history of smoking, and alcohol intake (information will be limited)
- Maternal obstetric history, including gravidity, parity, previous preterm deliveries, previous spontaneous abortions, or elective terminations, as available
- Medical conditions in pregnancy (such as course of MS; hypertension; diabetes; heart, thyroid, liver, kidney, and respiratory conditions (incl. asthma); and TORCH [Toxoplasmosis, Other (syphilis, varicella-zoster, parvovirus B19), Rubella, Cytomegalovirus (CMV), and Herpes] infections, as available)
- Medications from 6 months before conception through delivery, including other MS treatments during the study period; vaccinations; medications to treat diabetes, hypertension, and other diseases; teratogenic medications, and fertility treatment associated with this pregnancy
- Measures of health care services utilization in the 6 months before pregnancy: medications not related to the treatment of MS, health care encounters
- Sex of the infant

Data Sources

The data sources planned for this study are three health care claims databases in the US and in the national health care registries in Denmark:

- HealthCore Integrated Research Database℠ (HIRD) in the US
- Optum Dynamic Assessment of Pregnancies and Infants™ (DAPI) (formerly Systematic Tracking of Real Kids [STORK]) in the US
- IBM MarketScan Commercial Claims and Encounters Database (formerly Truven MarketScan Commercial Claims and Encounters Database MarketScan; hereafter, “MarketScan”) in the US
- The nationwide population-based registries and databases in Denmark

Study Size

The target for this study will be approximately 1,005 pregnancies exposed to ocrelizumab and 3,015 pregnancies in each of the two comparator cohorts, including women from the four proposed data sources. This size, assuming that 62% of pregnancies end in a live birth, and that 65% of pregnancies ending in live births are successfully linked to infants, will allow exclusion of a relative risk of 2.5 or greater, if the true relative risk is 1, for major malformations combined, which have a baseline prevalence of approximately 3% of live births.

Data Analysis

Comparisons will involve ocrelizumab-exposed pregnant women and two comparator cohorts: (1) primary comparison cohort—pregnancies in women with MS who have not been exposed to ocrelizumab (overall and in two strata—pregnancies exposed to non-ocrelizumab DMTs approved for the treatment of MS [subcohort 1a] and pregnancies not exposed to these DMTs [subcohort 1b]) and (2) secondary comparison cohort—pregnancies in women without MS who have not been exposed to ocrelizumab.

To control for confounding and channeling effect, exposed subjects will be matched in a variable 1:3 ratio (up to 1:3) with subjects from each of the two comparison cohorts (separately), using propensity scores.

Characteristics of the unmatched and matched cohorts, including the frequency of the outcomes, will be shown in tabular format. Balance in matching will be assessed by examining the distribution of variables in the cohorts and estimating standardized differences for each variable between the ocrelizumab-exposed and comparator cohorts. No statistical tests are planned for this comparison, but variables with standardized differences above 0.1 will be further evaluated and may lead to a re-evaluation of the propensity score estimation.

Unadjusted measures of outcome frequency will be estimated within the matched cohorts.

Measures of association will vary across outcomes and include incidence rate ratios and odds ratios. No adjustment is planned beyond matching.

Subgroup analyses will include strata of maternal age, calendar year, and others (depending on counts and data availability).

Sensitivity analyses will include modifications of the exposure window to start 26, 90, and 130 days before pregnancy (1, 3.5, and 5 mean half-lives of ocrelizumab, respectively), by trimester of pregnancy, and handling of covariates.

Results will be presented separately for each data source. Overall association results (e.g., odds ratios for major congenital malformations) will be summarized across data sources using meta-analytic techniques with random effects.

Milestones

Start of Study Observation

The study start date is the date when ocrelizumab is first available in one of the participating US data sources. The planned start date is Q2 2017, coinciding with the availability of ocrelizumab in the US.

End of Study Observation

The end of study observation is the date of the last observation included. The planned end-of-study-observation date is expected within the period Q2 2028-Q1 2029, depending on data availability within each data source.

5. AMENDMENTS AND UPDATES

None.

6. MILESTONES

After regulatory endorsement of the protocol, the study will be registered with the EU PAS Register (ENCePP, 2016). In each of the participating data sources, the study observation period (study period) will start with the first dispensing/prescription of ocrelizumab, which will vary by country (anticipated Q2 2017 in the US), and will end at varying times depending on the accrual of patients in the data source (anticipated within the period Q2 2028-Q2 2029), with anticipated study completion in Q2 2029 (Figure 1). Currently, data extraction from the first data source is anticipated in Q2 2028, with data extraction potentially occurring at different dates for each data source. This date depends on the drug uptake, which will in turn depend on the date of drug approval by health authorities, local reimbursement, availability of reimbursement codes (e.g., Healthcare Common Procedure Coding System procedure codes), and administrative procedures within the selected data sources and the data source lag time for the data release to research. For example, with a launch date of Q2 2017, data collection (i.e., extraction) from a data source would start in Q2 2028 assuming a 3- to 4-month lag in data availability (Figure 1).

Figure 1 Study Timeline

Study milestones are summarized in Table 1. The number of ocrelizumab-exposed pregnancies and live births will be monitored annually during the study observation period in all data sources to inform the study size, update the predicted study power, and determine the study data set creation date. Monitoring of drug uptake is planned to start in 2020. Completion of the final study report is planned for the end of Q2 2030, or earlier if study size is reached earlier. Study timelines will be confirmed at a later time based on

drug monitoring. Progress reports will be provided to the health authorities through scheduled regulatory safety reporting (periodic adverse drug experience reports [PADERs] and/or periodic benefit-risk evaluation reports [PBRERs], as agreed with health authorities).

Table 1 Study Milestones

Milestone	Planned Date
Registration of protocol in the EU PAS Register	Within 1 month of protocol approval
Start of study observation	Q2 2017
Monitoring of counts of ocrelizumab-exposed pregnancies and live births	2020
Monitoring of counts of ocrelizumab-exposed pregnancies and live births	2021
Monitoring of counts of ocrelizumab-exposed pregnancies and live births	2022
Monitoring of counts of ocrelizumab-exposed pregnancies and live births	2023
Monitoring of counts of ocrelizumab-exposed pregnancies and live births	2024
Monitoring of counts of ocrelizumab-exposed pregnancies and live births	2025
Monitoring of counts of ocrelizumab-exposed pregnancies and live births	2026
Monitoring of counts of ocrelizumab-exposed pregnancies and live births	2027
Monitoring of counts of ocrelizumab-exposed pregnancies and live births	2028
End of study observation	Q2 2028-Q2 2029
Start of data collection (EMA definition)[a]	Q2 2028
End of data collection (EMA definition)[b]	Q2 2029
Study completion	Q2 2029
Study progress report	According to PADER/PBRER schedule
Final report of study results (CSR)	Q2 2030 (within 1 year of end of the study)
Registration of the results in the EU PAS Register	Within 1 month of submission of final report to FDA

CSR = clinical study report; EU PAS Register = European Union electronic register of post-authorization studies; PADER = periodic adverse drug experience report; PBRER = periodic benefit-risk evaluation report.

Note: Contracts between the sponsor and research organization(s) and approvals by data protection, data custodian, ethics, and scientific and regulatory review bodies are pending. Timelines may be affected by approvals from these bodies, duration of contract reviews, and availability of data and staff at research institutions once contracts and approvals are finalized. If monitoring counts indicate that the target study size (Section 9.5) might be attained earlier than anticipated, the study will be launched earlier, allowing for earlier submission of the study report to the FDA.

[a] Start of data collection for secondary data use is “the date from which data extraction starts. Simple counts in a database to support the development of the study protocol, for example, to inform the sample size and statistical precision of the study, are not part of this definition” (EMA, 2017b).

[b] End of data collection for secondary data use is “the date from which the analytical data set is completely available” (EMA, 2017b).

7. RATIONALE AND BACKGROUND

Multiple sclerosis (MS) is a chronic inflammatory demyelinating disease of the central nervous system (CNS). Its etiology is not fully understood, but is thought to be a complex interaction of genetic susceptibility and environmental factors (Milo and Kahana, 2010). Multiple sclerosis affects approximately 2.3 million people worldwide, with the highest prevalence found in North America and Europe (140 and 108 patients with MS per 100,000 population, respectively) (MSIF, 2013). Overall, MS is more frequent in women than men, with a sex ratio of 2 to 3 women for every 1 man (Orton et al., 2006; Trojano et al., 2012), except in individuals with primary progressive form of the disease, which presents with similar prevalence by sex (NMSS, 2017b). Multiple sclerosis is commonly diagnosed during reproductive ages, between 20 to 50 years (NMSS, 2017d), with 30 years being the estimated average age of onset (MSIF, 2013).

About 85% of people with MS have relapsing-remitting MS, a form of MS with exacerbations or relapses during which new symptoms appear or previous symptoms worsen and periods of partial or complete remission. If left untreated, most patients who are diagnosed with relapsing-remitting MS will eventually progress to a chronic form characterized by progressive worsening of neurologic function over time with occasional relapses; which is categorized as secondary progressive MS. Approximately 15% of people with MS are diagnosed with primary progressive MS (PPMS), a form of MS with steadily worsening neurologic symptoms from the onset of disease and no clear relapses (Lublin et al., 2014; NMSS, 2017c).

Evidence suggests that inflammation is present in the CNS throughout all MS clinical courses, from relapsing-remitting MS to secondary progressive MS and in PPMS (Frischer et al., 2009). Differences between disease phenotypes are due to the differential contribution of each of the inflammatory and neurodegenerative processes to the pathophysiology of CNS damage over time (Frischer et al., 2009; Frischer et al., 2015).

OCREVUS® (ocrelizumab) was approved by the United States (US) Food and Drug Administration (FDA) on March 28, 2017, for the treatment of adult patients with relapsing forms of MS (RMS) and PPMS (Genentech Inc., 2017). Subsequently, OCREVUS® was approved in the EU, Switzerland, Australia, Canada, and other countries.

Ocrelizumab is a recombinant humanized monoclonal immunoglobulin G1 (IgG1) antibody that selectively targets CD20-expressing (CD20+) B cells. Two identical, randomized, active-controlled studies (OPERA I [Study WA21092] and OPERA II [Study WA21093]) have demonstrated superior efficacy outcomes versus interferon beta-1a in relapsing forms of multiple sclerosis (RMS) (Hauser et al., 2017); one randomized, placebo-controlled study (ORATORIO [Study WA25046]) demonstrated superior efficacy in PPMS versus placebo (Montalban et al., 2017). Results of these studies show that

depletion of CD20+ B cells leads to a significant impact on a broad range of clinical measures of disease, including disability progression, in addition to an impact on magnetic resonance imaging (MRI) outcomes related to disease progression and reflective of neural tissue loss, thus further supporting the hypothesis that B cells are central to the pathogenesis of both relapsing and primary progressive MS. Ocrelizumab has demonstrated a favorable safety profile in patients with RMS and PPMS (Hauser et al., 2017; Montalban et al., 2017). The proportion of patients with adverse events (AEs) was similar in patients treated with ocrelizumab and patients treated with interferon beta-1a (both 83.3%) or placebo (95.1% vs. 90.0%). The proportion of patients experiencing a serious adverse event was similar between ocrelizumab and the comparator groups—in RMS: 6.9%, ocrelizumab, and 8.7%, interferon beta-1a; in PPMS: 20.4%, ocrelizumab, and 22.2%, placebo.

7.1 STUDY BACKGROUND

Birth defects—structural (e.g., cleft lip/palate, heart defects, neural tube defects, abnormal limbs) and functional/developmental (e.g., sensory problems, metabolic disorders, nervous system problems, degenerative disorders)—affect about 3.0 per 100 live births in the United States (CDC, 2008) and are the leading cause of infant deaths (about 20% of all infant deaths) (Matthews et al., 2015). The European Surveillance of Congenital Anomalies (EUROCAT) estimated the prevalence of major congenital anomalies to be 23.9 per 1,000 live births from 2003-2007 (Dolk et al., 2010). It is uncertain whether MS has an impact on the risk of adverse pregnancy and infant outcomes. Some studies suggest there is little evidence that MS increases the risk of adverse pregnancy, delivery, or infant outcomes including perinatal mortality, congenital malformations, and delivery complications (Dahl et al., 2005; Houtchens, 2007; Mueller et al., 2002). Other studies suggest pregnancies in women with MS are associated with more frequent operative deliveries, decreased infant birth weight, and infants small for gestational age compared with women without MS (Dahl et al., 2008; Dahl et al., 2005; Kelly et al., 2009).

Ocrelizumab is a humanized monoclonal antibody of an immunoglobulin G1 subtype, and immunoglobulins are known to cross the placental barrier. B-cell levels in human neonates following maternal exposure to ocrelizumab have not been studied in clinical studies. There are no adequate and well-controlled data from studies in pregnant women; however, transient peripheral B-cell depletion and lymphocytopenia have been reported in infants born to mothers exposed to other anti-CD20 antibodies during pregnancy (Chakravarty et al., 2011; Klink et al., 2008).

It is not known whether ocrelizumab can affect pregnancy outcomes or infant outcomes in humans. However, based on pathophysiological considerations, ocrelizumab might theoretically affect pregnancy outcomes and infant outcomes in the following ways:

- By direct exposure of the fetus to ocrelizumab, which is assumed to occur after the 16th week of gestation as receptor-mediated transplacental transfer of IgG1; this is minimal during the first trimester of pregnancy (Palmeira et al., 2012; Simister, 2003).
- Indirectly due to known or unknown infections or infectious complications in the mother exposed to ocrelizumab during pregnancy, which may affect infants, where the infection may be associated with ocrelizumab exposure.
- Indirectly due to effects of ocrelizumab on the placenta.

Based on the average ocrelizumab terminal half-life of 26 days reported in the studies of MS (Genentech Inc., 2017), it is expected that ocrelizumab would be eliminated from the body approximately 4.5 months after the last administration. Considering the interpatient variability (the longest terminal half-life recorded in women was 53 days) and the absence of placental transfer of immunoglobulins during the first trimester of pregnancy, it is recommended that women of childbearing potential should use contraception while receiving ocrelizumab and for 6 months after the last infusion of ocrelizumab. However, women may get pregnant during this period due to noncompliance with contraception or failure of contraceptive methods.

Clinical studies of the effects on pregnancy and infants associated with the use of ocrelizumab during pregnancy, during lactation, and/or before conception have not been performed, and experience from ocrelizumab clinical trials is very limited. As of 31 January 2017, 25 pregnancies were reported during ocrelizumab trials in MS. Of these 25 pregnancies, 7 were electively terminated, and no abnormalities were found in the embryos or products of conception; 1 pregnancy ended in stillbirth at approximately 7-8 months of gestation; 2 live, preterm births had abnormal findings (benign nasopharyngeal neoplasm, jaundice, respiratory distress, and low birth weight for one infant and temperature instability, feeding difficulties, bradycardia, respiratory distress, and anemia for the other); 11 pregnancies ended in a live, full-term birth; and 4 pregnancies were still ongoing at the time of analysis (Vukusic et al., 2017).

In an embryo-fetal development study in cynomolgus monkeys, there was no evidence of teratogenicity or embryotoxicity (Study 04-1272-1342). Also in monkeys, using doses similar to or larger than those used clinically on a mg/kg basis, increased perinatal mortality (in some cases associated with bacterial infections), depletion of circulating B cells, and toxicity to kidneys (glomerulopathy and inflammation), bone marrow (lymphoid follicle formation), and testis (reduced weight) were seen in the offspring in the absence of toxicity in the mother (Genentech Inc., 2017). Transient peripheral B cell depletion and lymphocytopenia have been reported in infants of women treated during pregnancy with other drugs with the same mechanism of action. Ocrelizumab was excreted in the milk of

treated monkeys. The effects on the infant of ocrelizumab exposure through lactation are not known (Genentech Inc., 2017).

Roche proposes a non-interventional multidatabase post-marketing safety study to assess pregnancy-related safety data from women with MS exposed to ocrelizumab. Research partners (HealthCore, Inc.; Optum; RTI Health Solutions [RTI-HS]; and [redacted]) will conduct the work in each data source, and a coordinating center (RTI-HS) will lead/oversee the development of study documents, coordinate research activities, and pool results from individual data sources. This study will be conducted to fulfill part of the FDA post-marketing requirements (PMR 3194-4) for approval of ocrelizumab in the US. The study is listed as additional pharmacovigilance activity to address the missing information on drug use during pregnancy and lactation in the proposed European Union Risk Management Plan.

8. RESEARCH QUESTION AND OBJECTIVES

To assess and characterize pregnancy and infant outcomes of women with MS exposed to ocrelizumab during the 6 months before the estimated date of conception or at any time during pregnancy.

The objectives are as follows:

- To estimate the frequency of selected adverse pregnancy outcomes in women with MS exposed to ocrelizumab during the defined exposure window (i.e., spontaneous abortions, fetal death/stillbirths, elective abortions, preterm births, C-sections, and urinary and other infections in pregnancy)
- To estimate the frequency of selected adverse fetal/neonatal/infant outcomes at birth and up to the first year of life of infants from pregnancies in women with MS exposed to ocrelizumab—i.e., major congenital malformations, small for gestational age, adverse effects on immune system development (e.g., severe infectious disease in the first year of life)
- To compare the frequency of each safety event of interest between ocrelizumab-exposed pregnant women with MS and two comparison cohorts: (1) primary comparison cohort—pregnancies in women with MS who have not been exposed to ocrelizumab (overall and in two strata—pregnancies exposed to any non-ocrelizumab disease-modifying therapies [DMTs] approved for the treatment of MS or any new DMT approved during the study period [subcohort 1a], and pregnancies not exposed to these DMTs [subcohort 1b]) and (2) secondary comparison cohort—pregnancies in women without MS who have not been exposed to ocrelizumab.

9. RESEARCH METHODS

In addition to a prospective product-specific pregnancy registry (PMR 3194-3, protocol WA40063) in the US, Germany, and possibly other countries, Roche proposes this study, in which secondary data sources will be used to ensure collection of sufficient

data to reach scientifically valid conclusions on the safety of ocrelizumab use before conception and during pregnancy as soon as possible after marketing authorization. Roche is aware that Marketing Authorization Holders of other MS treatments (e.g., fingolimod, natalizumab, alemtuzumab, and teriflunomide) have attempted to generate pregnancy safety data by establishing drug-specific pregnancy registries. It has been noted by both the FDA and the EMA that product-specific pregnancy registries very often fail to provide clinically meaningful information in a timely manner, because of inadequate enrollment, loss to follow-up, selection bias, limited statistical power, and lack of suitable comparator populations (Gelperin et al., 2017; Sahin, 2014). A recent analysis of pregnancy registries found that among products that had a stated target enrollment in the pregnancy registry protocol (22 of 59 registries [37%]), only 14% (3 of 22) achieved target enrollment (Sahin, 2014). In the case of MS, two therapy-specific pregnancy exposure registries had to be terminated early because of low enrollment, and other ongoing pregnancy registries have experienced delays due to recruitment challenges (Anthony and Krueger, 2016). Given these known challenges and limitations in obtaining quality pregnancy exposure data for newly launched drugs, Roche aims to obtain more timely information on the safety of the use of ocrelizumab before conception and during pregnancy by utilizing health care databases from the US and Europe. Analyses of existing databases with prospectively collected data have been proven useful to evaluate many drug safety questions, including questions on drug safety in pregnancy (Broms et al., 2014; Charlton et al., 2016; Hviid et al., 2013; Johansson et al., 2015; Li et al., 2014; Mines et al., 2014; Molgaard-Nielsen et al., 2016). Using existing databases minimizes the risk of self-selection into the study that is inherent to pregnancy exposure registries. Self-selection into a study may affect generalizability in that volunteers may be different from the average patient who is treated in routine practice. Additionally, the validity of relative risk estimates may be affected if subsets of women enroll late in registries (after the at-risk period for early pregnancy outcomes such as spontaneous abortions) or if the likelihood that women will enroll depends on their risk for outcomes such as congenital malformations.

In the past, pregnancy exposure registries were considered the first line of evidence because they are able to enroll women as soon as the drug is in the market. Standardization of protocols, methods development, and international collaborations are now allowing the creation of nested pregnancy cohorts within health care databases that are pooled, resulting in larger sample sizes in a shorter time.

Because the number of eligible subjects may be small in any given database, this non-interventional post-marketing safety study will integrate results from multiple data sources that prospectively collect relevant pregnancy data. Mother-child linked data will be extracted from existing health care claims databases in the US and population-based patient registries in Denmark. These two countries have a high prevalence of MS (> 100 people with MS per 100,000 population) (MSIF, 2013) and are generally quick in the uptake of new medications. Specifically, in Denmark, the prevalence of MS was 250 per

100,000 persons, with a female to male ratio of 2 to 1 in 2017 (personal communication from [redacted], [redacted], July 2017). Suitability of the data sources will be confirmed once reimbursement status is known. Furthermore, drug uptake will be monitored approximately every 2 years to assess when the target study size is reached.

9.1 STUDY DESIGN

This is a matched cohort study using multiple sources of prospectively collected secondary data. Selected outcomes will be validated in the data sources in which this is possible. Study subjects will include pregnant women and their children born during the study period.

The following cohorts will be assembled:

- Pregnancies in women with MS and exposure to ocrelizumab initiating or continuing within 6 months before conception and/or at any time during pregnancy, and the children born to these pregnancies (exposed cohort). For analyses on congenital malformations, the exposure period will be 6 months before conception or any time during the first trimester of pregnancy. For analyses on spontaneous abortions and elective terminations, the exposure period will be 6 months before conception or any time during the first trimester if prior to the end of pregnancy.
- Pregnancies in women with MS but no ocrelizumab exposure in the 6 months before conception or at any time during pregnancy who are matched to ocrelizumab-exposed pregnancies, and the children born to these pregnancies during the study period (primary comparison cohort). For analyses on congenital malformations and spontaneous abortions and elective terminations, the exposure windows will be the same as those for ocrelizumab-exposed pregnancies. Results will be presented for the overall cohort and for two mutually exclusive strata:
  - Pregnancies in women with MS exposed to any non-ocrelizumab DMTs approved for the treatment of MS or to any new DMTs approved during the study period (primary comparison subcohort 1a)
  - Pregnancies in women with MS not exposed to DMTs approved for the treatment of MS (primary comparison subcohort 1b)
- Pregnancies in women without MS or ocrelizumab use who are matched to ocrelizumab-exposed pregnancies, and the children born to these pregnancies during the study period (secondary comparison cohort).

9.2 SETTING

The data sources for this study include health care claims from the US and country-level registries from Denmark, including the Danish MS Registry.

A common core protocol will be adapted to each data source so that methods are as homogeneous as possible across data sources but also tailored to the specific characteristics of each source.

9.2.1 Eligibility Criteria

In each data source, the study will include all matched pregnancies in women with a pregnancy in the study period with continuous enrollment with pharmacy benefits in the 6 months before the estimated beginning of pregnancy and throughout the pregnancy and the children born to these pregnancies during the study period. For pregnancy outcomes, linkage to infants is not required (i.e., pregnancies not linked to an infant will be retained). For infant outcomes, linkage between the mother and infant is required.

Matching will be based on an exposure propensity score, which is described in Section 9.7.4.

A woman with MS and two recorded pregnancies in the study period who received ocrelizumab in one pregnancy and another drug to treat MS in the other pregnancy will contribute one pregnancy to each of two cohorts.

Women who used both ocrelizumab and a non-ocrelizumab DMT defining comparison subcohort 1a in the exposure window will be excluded from the comparative analyses, because their information will not be useful to understand whether any effect seen is associated with ocrelizumab or a comparator drug. Counts of pregnancies and counts of outcomes from these pregnancies, by drug combinations, will be presented in the study report. This situation is expected to occur in a small number of pregnancies. Women will be excluded from the study if they received, from 6 months before pregnancy or during pregnancy, rituximab, a monoclonal antibody with the same mechanism of action as ocrelizumab that is used off-label for the treatment of MS (Salzer et al., 2016); ofatumumab; or other drugs with the same mechanism of action that may become available during the study period.

Eligibility criteria for children are inclusion of their mother in one the three study cohorts (i.e., successful mother-child linkage is required) and continuous enrollment covering outpatient care and hospitalizations (disenrollment or gaps in enrollment will stop follow-up, as specified in Section 9.2.3, Follow-up).

9.2.2 Study Period

The study observation period (study period) will start with the first dispensing/prescription of ocrelizumab in each of the participating data sources, which will vary by country. Accrual and follow-up of pregnant women and their infants will start with the first ocrelizumab prescription following approval on March 28, 2017 in the US and on January 8, 2018 in the European Union, and will end at the latest date with data collected from each data source (expected in 2028). The latest date with data may vary

across data sources, because the study will attempt to include the largest possible number of pregnancies from each data source (see Section 6, Figure 1).

9.2.3 Follow-up

Follow-up of women will start at the estimated beginning of pregnancy and will finish at the end of pregnancy; follow-up of infants will start at birth and finish at 1 year of age. For mothers and infants, follow-up will finish at the earliest of death, disenrollment from the data source, or end of the study period. Information from before cohort entry will be used to assess baseline characteristics. For each outcome of interest that can occur multiple times, follow-up for that outcome will stop at its first occurrence (e.g., urinary tract infections in pregnancy, infections requiring hospitalization in pregnancy). Follow-up will continue for other outcomes.

9.3 VARIABLES

9.3.1 Exposure

The exposure of interest is ocrelizumab; other DMTs approved for MS, which will be the basis for forming comparison subcohort 1a, are listed in Table 2 .Other MS DMTs that may become available during the study period will be considered for inclusion in this list (except those listed for exclusion [Section 9.2.1]). Such new drugs will be documented in the statistical analysis plan (SAP).

The exposure window will include the time from 6 months before conception until the end of pregnancy for end-of-pregnancy and infant outcomes, until the end of the first trimester for analyses on congenital malformations, until the end of the first trimester or end of pregnancy if prior to end of first trimester for analyses on spontaneous abortions or elective terminations, or to the date of occurrence of outcomes that may occur before the end of pregnancy (e.g., infections during pregnancy). The interval of 6 months before conception takes into account ocrelizumab half-life, as described in Section 7, Rationale and Background, and is based on the prescribing information in the US at the time of approval (March 2017). For other drugs, the exposure window will be as that for ocrelizumab, despite varying half-lives, taking into consideration that prescriptions may cover a few months. Medications used before 6 months before conception will not play a role in determining entry into any comparison subcohorts in the primary analysis.

Exposure will be defined based on records of one or more dispensed prescriptions (treatment initiation or continuation) with dispensing dates within the noted exposure windows. Additional details are provided in Table 7 in Section 9.7.1, Study Cohorts.

The sources for ascertainment of medication use will be pharmacy claims and product-specific administration procedure codes in claim databases and treatment information in the MS Registry in Denmark (in Denmark, pharmacies do not dispense MS treatments). The timing of medication use relative to the beginning of pregnancy will be estimated, if needed, using previously described methods (Margulis et al., 2015).

Table 2 Ocrelizumab and Other Multiple Sclerosis Disease-Modifying Therapies

ATC Code[a]	Substance[b]	Route and Frequency of Administration[b]	Duration of Contraception After Last Course of Treatment[c]
L04AA34	Alemtuzumab	IV infusion in two blocks (5 and 3 days) 12 months apart	4 months
L01BB04	Cladribine	Oral with 1 week of treatment at the start of treatment, another week of treatment 1 month later, and a second cycle in the second year	6 months
L04AC01	Daclizumab[d]	SC once a month	Not specified
N07XX09	Dimethyl fumarate	Oral twice daily	Not specified
L04AA27	Fingolimod	Oral daily	2 months
L03AX13	Glatiramer acetate	SC every day or 3 times a week	Not specified
L03AB07	Interferon beta-1a	IM once a week or SC 3 times a week	Not specified
L03AB08	Interferon beta-1b	SC every other day	Not specified
L01DB07	Mitoxantrone	IV infusion every 3 months	Not specified
L04AA23	Natalizumab	IV infusion every 28 days	Not specified
L04AA36	Ocrelizumab	IV infusion every 2 weeks as a starting dose and every 6 months later	6 months
L03AB13	Peginterferon beta-1a	SC every 14 days	Not specified
L04AA31	Teriflunomide	Oral daily	Until it is verified that plasma concentrations of teriflunomide are less than 0.02 mg/L (0.02 mcg/mL)

ATC = Anatomical Therapeutic Chemical; IM = intramuscular; IV = intravenous; NMSS = National Multiple Sclerosis Society; SC = subcutaneous; WHO = World Health Organization.

[a] WHO (2017).

[b] NMSS (2017a).

[c] Sources: Drugs@FDA: FDA Approved Drug Products https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm; European product information at https://www.ema.europa.eu/en/medicines.

[d] Withdrawn as of March 2, 2018.

9.3.2 Outcomes

Outcomes will be defined homogeneously across the data sources to the fullest extent possible based on the available information. The study outcomes are listed below. For diagnoses that are less commonly studied in health care databases, references supporting the feasibility of the approach are provided. When outcomes definitions vary across data sources, explanatory text is provided.

(1) Pregnancy outcomes and pregnancy complications

- Spontaneous abortion [pregnancy loss at < 20 completed weeks (ACOG, 2009)]. Abortion events with codes for ectopic pregnancy will not be considered events of spontaneous abortion.
- Elective termination: the reason for termination (e.g., therapeutic abortion, abnormal findings in fetus in prenatal tests, ectopic pregnancy) will be ascertained to the extent to which data are available.
- Fetal death/stillbirth (≥ 20 completed weeks). The standard definition of stillbirth in the Danish Medical Birth Registry uses 22 weeks of gestational age as the threshold, but it will be adapted to the preferred definition in this study; earlier fetal demises will be considered spontaneous abortions.
- Preterm delivery (live birth at < 37 completed weeks).
- C-section (emergency or elective cesarean sections).
- Urinary tract infection during pregnancy (acute cystitis or asymptomatic bacteriuria in data from ambulatory care, and dispensing of appropriate antibiotics). Danish registries will not capture mild episodes treated only in the primary care setting; for these occurrences, the outcome will be based on dispensed antibiotics.
- Infections requiring hospitalization during pregnancy (hospitalizations with main discharge diagnosis being an infectious disease or chorioamnionitis in any position). Chorioamnionitis will be reported separately.

(2) Fetal/neonatal/infant outcomes

- Major congenital malformations (MCMs). Definitions and potential groupings will be based on guidelines from the European Surveillance of Congenital Anomalies program (EUROCAT, 2012). In addition to MCMs diagnosed in liveborn infants, information on MCMs from spontaneous abortions, stillborns, and elective terminations will also be used, if available. Patent foramen ovale, persistent ductus arteriosus, or undescended testes will not be included among MCMs; patent foramen ovale, persistent ductus arteriosus, and possibly other cardiac malformations will be presented separately in a category named “unspecified cardiac defects.” The rationale is that these malformations are often physiologically expected in preterm births or are the result of improved technology, with little clinical significance for a large proportion of cases.

- Specific categories of MCMs (e.g., cardiovascular) and specific malformations will be explored depending on the number of events observed (e.g., hypospadias, cleft lip with or without cleft palate, and cardiac malformations [or subtypes]). MCMs are not expected through direct effects of the drug, as transplacental transfer of IgG1 is minimal before the 16th week of gestation (Palmeira et al., 2012; Simister, 2003).
- Minor malformations will be examined to the extent available. Codes to define minor malformations will be based on guidelines from EUROCAT (2014). Underrecording is expected in all data sources. A category for unspecified cardiac defects will be included.
- Small for gestational age (defined by birth weight < 10th percentile of the sex- and gestational-age-specific birth weight for data sources for which birth weight data are available, or diagnosis codes for small for gestational age). A secondary algorithm will be based on the fifth percentile.
- Infant growth (length, weight, and head circumference), including measurements at birth and during follow-up, will be explored where available. Danish registries are anticipated to contain some information on head circumference at birth.
- Infant development will be explored to the extent available.
- The following adverse effects on the immune system in the first year:
  - Hospitalizations due to infectious diseases, stratified by neonatal infections (within 28 days of birth) and later infections. The rationale for this stratification is that fever in neonates generally triggers a much more intensive sepsis workup.
  - Any cancer, including leukemia.
  - Vaccine-preventable diseases and vaccine-associated poliomyelitis in the first year of life: composite outcome of hepatitis B, whooping cough, tetanus, diphtheria, rotavirus diarrhea, invasive *Haemophilus influenza* b disease, invasive pneumococcal disease, poliomyelitis and vaccine-associated poliomyelitis. Included in this composite outcome are infectious diseases for which children are typically immunized before 1 year of age. Among all possible occurrences of these conditions, an event will be considered to be a study outcome only if the event occurs between the age of first immunization (as listed in the country-specific immunization guidelines and recommendations) for the condition and 1 year of age of the infant.

Outcomes will be ascertained from diagnosis codes and procedure codes in US and Danish data sources in maternal and infant inpatient and outpatient records; dispensed drugs will be used in both types of data sources for ascertainment of treatment of urinary infections. Outcome definitions will be expanded in the SAP.

The ability of data sources to support validation efforts was considered in the feasibility assessment. Of the proposed data sources, outcome validation is feasible in HealthCore and Optum DAPI data. In MarketScan data, outcomes will be identified using algorithms validated in other US claims data sources. The approach for validation may include review of electronic medical records, medical chart abstraction, and review of claims, with estimation of the positive predictive values of outcome codes or algorithms (point

estimate and 95% confidence interval). Cases of selected outcomes in the matched cohorts might be validated, subject to availability of medical charts, if the chosen validation method requires review of medical charts. The selection of outcomes for validation will be informed by experts in the subject matter and in the data sources so that efforts are focused on the outcomes more susceptible to misclassification. The description of this process will be incorporated in the Data Validation Plan, along with the strategy for estimating the target number of events for validation and the strategy for sampling events. These strategies will be determined in collaboration with the research partners to ensure that nuances of each data source are taken into consideration.

As a part of the outcome validation process, and if no validated algorithms for preterm birth using ICD-10 codes are available before the conduct of the analysis, three electronic algorithms for ascertaining preterm birth will be assessed; the prevalence of preterm birth will be estimated using these three algorithms in US data sources and one in Danish data (Appendix 3). Similarly, if no validated algorithms for small for gestational age (SGA) using ICD-10 codes are available before conduct of the analyses, three electronic algorithms will be assessed in US data sources, and the prevalence of SGA will be estimated using these three algorithms in US data sources (Appendix 3). In Danish data, prevalence of SGA will be estimated using primary and secondary algorithms (Appendix 3). Validation efforts conducted in US data sources will determine which algorithms for preterm birth and SGA will be selected for use in subsequent analyses in US data sources. For selected outcomes, if validated algorithms are not available for implementation in MarketScan, the number of outcomes observed in MarketScan will be placed into context by estimating the number of outcomes that would be expected had the algorithm performed with the same accuracy as in Optum DAPI and HIRD.

A side-by-side comparison of estimated prevalences from the various study data sources (without statistical tests) is planned to help the research team interpret similarities and differences across data sources.

To the extent possible, validated algorithms (validated within this study or published) will be used to identify outcomes. As the transition from ICD-9 to ICD-10 codes happened relatively recently in US claims data (October 1, 2015), limited validation work has been published at this point. Algorithms to define key study endpoints are presented in Appendix 3; it was not possible to define these endpoints with validated ICD-10 codes. Validated algorithms for other endpoints will be sought and specified in the SAP. If validated algorithms using ICD-10 or ICD-10-CM codes in US claims data with positive predictive value (PPV) (point estimate) ≥ 80% become available before the conduct of this study, those validated algorithms will be considered for use instead of the ones proposed in Appendix 3.

Details will be defined in a DVP to be developed closer to the study start date in consultation with clinicians with appropriate expertise within each collaborating institution

(e.g., teratology or pediatrics with expertise in birth defects) who will support the clinical review process.

9.3.3 Covariates

Patient and pregnancy characteristics of interest are listed in Table 3. These variables were selected to provide a general description of the health and characteristics of the study cohorts and to explore potential confounding. The level of detail and completeness will vary across data sources. In claims data, the period for ascertaining baseline characteristics will be limited by the period of enrollment in the health care plan. Obesity, smoking habits, alcohol abuse and drug abuse may be very incompletely captured in all data sources. Their inclusion in analyses will be decided based on the information available (e.g., percentage of missing values).

Table 3 Covariates

Covariate	Comments	Availability in Data Sources			
		HealthCore	MarketScan	Optum	Danish National Registries
Maternal characteristics at conception					
Age	In years	X	X	X	X
Calendar year	In years	X	X	X	X
Maximum education attained	In categories (to be specified)	NA	NA	NA	X
Obesity	Diagnosis codes and through proxies (treatment); likely to be incompletely captured in all sources	X	X	X	X
Smoking habits	Defined through proxies (treatment) or diagnosis codes, likely underrecorded in all sources	X	X	X	X
Alcohol abuse	Defined through diagnosis codes and proxies (treatment), likely underrecorded in all sources	X	X	X	X
Drug abuse	Defined through diagnosis codes and proxies (treatment)	X	X	X	X
MS type	E.g., RRMS. ICD-10 and ICD-10-CM do not have entries for MS types In Denmark, information will be extracted from the MS Registry	NA	NA	NA	X
Duration of MS	In claims, time since first diagnosis code (risk of underascertainment due to left censoring, or restricted look-back period). In Denmark, based on year of onset as recorded in the MS Registry.	X	X	X	X
Number of MS relapses in 6 months previous to conception	In claims, relapses will be identified using the same algorithm across data sources. Published algorithms with their validation results are presented in Table 4. The selected algorithm will be specified in the SAP. In Denmark, information will be extracted from the MS Registry.	X	X	X	X

Covariate	Comments	Availability in Data Sources			
		HealthCore	MarketScan	Optum	Danish National Registries
DMTs approved for MS received in the 12 months before the exposure window	From list in Table 2 Dispensed prescriptions and product-specific administration procedure codes In Denmark, information will be extracted from the MS Registry.	X	X	X	X
Other drugs used for MS received before conception (e.g., methotrexate, azathioprine, cyclophosphamide, rituximab)	Dispensed prescriptions and product-specific administration procedure codes	X	X	X	X
Drugs to treat MS symptoms (e.g., amantadine, baclofen)	From dispensed prescriptions	X	X	X	X
Hypertension	Defined through diagnoses, or use of medications in previous 12 months, not including gestational hypertension	X	X	X	X
Diabetes	Defined through diagnoses or use of medication in previous 12 months, not including gestational diabetes	X	X	X	X
Thyroid disease	Defined through diagnosis codes and applicable medications	X	X	X	X
Heart disease	Defined through diagnosis codes and applicable medications	X	X	X	X
Respiratory disease, incl. asthma	Defined through diagnosis codes and applicable medications	X	X	X	X
Liver disease	Defined through diagnosis codes and applicable medications	X	X	X	X
Kidney disease	Defined through diagnosis codes and applicable medications and procedures	X	X	X	X

Covariate	Comments	Availability in Data Sources			
		HealthCore	MarketScan	Optum	Danish National Registries
Non-MS chronic neurologic disease	Defined through diagnosis codes and applicable medications	X	X	X	X
Malignancies except nonmelanoma skin cancer and in situ cancer	Defined through diagnosis codes and applicable medications	X	X	X	X
Anxiety/depression	Defined through diagnoses or use of medication in previous 12 months	X	X	X	X
Severe mental health conditions	Hospitalizations or institutionalization, not including depression	X	X	X	X
Gravidity	Defined through diagnosis codes in claims data (limited by duration of enrollment before cohort entry) and in Medical Birth Registry in Denmark. Spontaneous abortions and terminations may be underrecorded	X	X	X	X
Parity	As number of deliveries or C-sections. Limitations as above. In Denmark, information will be extracted from the Danish Medical Birth Registry.	X	X	X	X
Spontaneous abortions in previous pregnancies	Defined through diagnosis codes. Likely underrecorded in all sources. In Denmark, information will be extracted from the Danish National Patient Registry (DNPR).	X	X	X	X
Pregnancy terminations in previous pregnancies	Defined through diagnosis codes. Likely underrecorded in all sources. In Denmark, information will be extracted from the Danish National Patient Registry (DNPR).	X	X	X	X

Covariate	Comments	Availability in Data Sources			
		HealthCore	MarketScan	Optum	Danish National Registries
Health care utilization in the 6 months before conception					
Psychotropic medications	Defined through dispensed prescriptions	X	X	X	X
Medications not used to treat MS or its symptoms	Defined through dispensed prescriptions	X	X	X	X
Number of medical encounters	Count of encounters with different date	X	X	X	X
Descriptive data for this pregnancy					
Multiple pregnancy	Defined through diagnosis codes. In Denmark, information will be extracted from the Medical Birth registry	X	X	X	X
Gestational diabetes	Defined through diagnosis codes. In Denmark, information will be extracted from the Danish National Patient Registry (DNPR)	X	X	X	X
Preeclampsia/ eclampsia	Defined through diagnosis codes. In Denmark, information will be extracted from the Danish National Patient Registry (DNPR).	X	X	X	X
Number of MS relapses during pregnancy	Please see definition of proxy for relapses provided earlier in this table and in Table 4. In Denmark, information will be extracted from the MS Registry.	X	X	X	X
Sex of the infant	Administrative records	X	X	X	X
Use of oral contraceptives around conception	Defined through dispensed prescriptions in the 3 months before pregnancy and the first month of pregnancy	X	X	X	X
TORCH infections	Defined through diagnosis codes for toxoplasmosis, syphilis, varicella-zoster, parvovirus B19, rubella, cytomegalovirus, and herpes infections during pregnancy (Stegmann and Carey, 2002). Only infections that triggered contact with health care will be identified.	X	X	X	X

Covariate	Comments	Availability in Data Sources			
		HealthCore	MarketScan	Optum	Danish National Registries
Teratogenic medications	Prescriptions dispensed from 6 months before conception to end of first trimester and end of third trimester of pregnancy for drugs in list of teratogenic medications in Eltonsy et al. (2016)	X	X	X	X
Vaccines	Dispensed prescriptions (vaccines offered at work, e.g., flu, will be underrecorded)	X	X	X	X

DMT = disease-modifying therapies; ICD-10 = International Statistical Classification of Diseases and Related Health Problems, 10th Revision; ICD-10-CM = International Classification of Diseases, 10th Revision, Clinical Modification; MCM = major congenital malformation; MS = multiple sclerosis; NA = not available; RRMS = relapsing-remitting multiple sclerosis; TORCH = Toxoplasmosis, Other (syphilis, varicella-zoster, parvovirus B19), Rubella, Cytomegalovirus (CMV), and Herpes infections; X = information available.

Table 4 Published Algorithms to Identify Relapses of Multiple Sclerosis in Claims Data

Reference, Data Type, and Country	Summary of Algorithm	Validation Results
Chastek et al. (2010) Health care claims from the US	Relapses were identified as follows: ▪ Hospitalizations with a primary diagnosis of MS at any time during hospitalization, OR ▪ A corticosteroid claim following an outpatient visit with a code for MS in the primary or secondary position	Validation of cases was conducted against medical charts for 300 patients. ▪ PPV = 67.3% ▪ NPV = 70.0%
Wang et al. (2015) Health care claims from Optum Research Database, US	Relapses were identified as follows: Algorithm 1: ▪ High dose (≥ 500 mg per day) of oral prednisone, prednisolone, or methylprednisolone for up to 15 days; OR ▪ ATCH ≥ 80 U per day for 5 days or more; OR ▪ Any IV methylprednisolone (regardless of dosage) for up to 15 days Algorithm 2: ▪ High dose (≥ 500 mg per day) of oral prednisone, prednisolone, or methylprednisolone for up to 15 days; OR ▪ ATCH ≥ 80 U per day for 5 days or more; OR ▪ Hospitalization episode associated with MS (ICD-9 340)	No direct validation, but authors report that these algorithms identified relapses in 1.4% and 2.1% of patients treated with dimethyl fumarate for 6 or 12 months, respectively, while previous reports were around 6% in patients 6 and 12 months after initiation of dimethyl fumarate.

ACTH = adrenocorticotropic hormone; ICD-9-CM = International Classification of Diseases, 9th Revision, Clinical Modification; MS = multiple sclerosis; NPV = negative predictive value; PPV = positive predictive value.

9.4 DATA SOURCES

This is a multisource study involving population-based patient registries and health care databases that prospectively collect data in the US and Europe. The specific data sources were determined in a feasibility study conducted to identify the most appropriate sources for this study. Availability of mother-infant linked data, birth certificates, and other data, and ability to support partial outcome validation was assessed in the feasibility study (Anthony et al., 2016). The assessment evaluated specific information on the totality of data available in the databases and operational aspects that could confirm the ability to identify specific cohorts, pregnancies, medication exposure, and outcomes of interest for the study of ocrelizumab in pregnancy.

The study is to be conducted in three health care claims databases in the US and in the national health care registries in Denmark:

- HealthCore Integrated Research Database℠ (HIRD) in the US. Analyses will be conducted by HealthCore.
- Optum Dynamic Assessment of Pregnancies and Infants™ (DAPI) in the US. Analyses will be conducted by Optum.
- IBM MarketScan Commercial Claims and Encounters Database (formerly Truven MarketScan Commercial Claims and Encounters Database) in the US. Analyses will be conducted by RTI-HS.
- The National Health Databases in Denmark and the Danish MS Registry. Analyses will be conducted by [redacted]

The top three choices for US claims databases were HIRD, DAPI, and MarketScan, based on the availability of the required information for the study on exposures, patient characteristics, and outcomes; the possibility to link mother and infant data; and the largest numbers of patients with MS among the databases evaluated. HIRD and DAPI both have access to medical records for at least a portion of their participants, so outcomes can be validated. While MarketScan might provide the largest number of ocrelizumab-exposed pregnancies, outcome validation against medical records cannot be done, but accuracy of results could be evaluated by comparing results from data sources with outcome validation.

The Danish National Health Databases and MS Registry are nationwide registers that offer the possibility of maternal data to be linked with birth defect registers, in which completeness of ascertainment and quality of diagnosis has been shown to be high. In addition, linking health, civil, and administrative data enriches the amount of data available beyond individual data sources, and follow-up of patients is virtually lifelong.

Key characteristics of the study data sources are described in Appendix 4.

9.4.1 HealthCore Integrated Research Database℠

HealthCore is a wholly owned subsidiary of Anthem, Inc., which is the largest health benefits company in terms of medical membership in the US. Anthem is an independent licensee of the Blue Cross and Blue Shield Association. HealthCore is the health services research entity for Anthem that integrates the public health, pharmacoepidemiologic, health outcomes, and pharmacoeconomic concerns of these companies and their clients to conduct outcomes analyses. HealthCore maintains the HIRD for use in health services research.

The HIRD is a broad, clinically rich, and geographically diverse spectrum of longitudinal medical and pharmacy claims data from health plan members in the Northeastern, Mid-Atlantic, Southeastern, Midwestern, Central, and Western regions of the US. The database represents claims information from one of the largest commercially insured populations in the US. Patient enrollment data, inpatient and outpatient medical care, outpatient prescription drug use, outpatient laboratory test results, and health care charges may be tracked for each patient in the database dating back to 2006 (HealthCore, 2017).

In the HIRD, diagnoses and procedures will be identified by the following types of codes for both outpatient visits and inpatient stays: *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM); *International Classification of Diseases, 10th Revision, Clinical Modification* (ICD-10-CM); Current Procedural Terminology; and Healthcare Common Procedure Coding System. Drug claims are captured by National Drug Codes, which can then be translated to broader categories of coding such as Generic Product Identifier codes. Information on physician specialty is also available in the HIRD. In addition, the HealthCore Integrated Research Environment (HIRE) has the ability to link the claims data in the HIRD to complementary data sources, including inpatient and outpatient medical records for the health plan members represented in the HIRD; identify and contact providers and members for survey research through vendor relationships; and link data to national vital records, such as the National Death Index (NDI), for date and cause of death (CDC, 2016b). Death records are added to the NDI file annually, approximately 12 months after the end of the calendar year. Cause-of-death codes can be obtained using the NDI Plus service.

For women with a recorded delivery, linked infants who are captured in the HIRD can be identified by requiring that the infant share the mother's subscriber identification number and have a date of birth within 30 days of the recorded delivery date. The performance of alternative linking strategies in the event that subscriber identification numbers are unavailable can be explored. In past studies involving linkage of mothers and their infants within the HIRD, approximately 70% to 75% of completed pregnancies could be connected to a qualifying infant. When the infant is not identifiable, it is likely that the infant was covered by the insurance plan of the other parent.

9.4.2 Optum Dynamic Assessment of Pregnancies and Infants™ (DAPI)

Within the Optum data source, DAPI data will be used to identify pregnancies and link the health care data of mothers with that of their infants within a large health care claims database that covers members of a large health insurer affiliated with Optum. The database contains medical claims and enrollment data dating back to 1994. Patient and physician data are linked to pharmacy and medical claims, medical record data, socioeconomic measures, and clinical laboratory results. Estimates from recent years show that the size of the source database provides approximately 95,000 pregnancies each year, on average, of which, 84% resulted in live births, and 86% of live births have the mother linked to the infant. Data of mothers and infants is linked through a family identifier and by matching the dates of delivery and infant's birth. The fraction of identified deliveries that cannot be matched to an infant is likely due to the infant being carried under a different health insurance from the mother, in other instances, this is due to pregnancies that did not end in a delivery. Approximately 30% to 40% of identified pregnancies cannot be linked to an infant (Wyszynski et al., 2016).

Because the linkage is made within an identifiable health insurance database affiliated with Optum, Optum can (with appropriate approvals) access medical records for mothers or infants to ascertain covariate information or to confirm outcomes.

9.4.3 IBM MarketScan Commercial Claims and Encounters Database

The MarketScan claims database is a large convenience sample with more than 41.1 million covered lives in the most recent full data year. The database encompasses employees, their spouses, and their dependents covered by employer-sponsored private health insurance. More than 300 employers and 40 contributing health plans throughout the US are represented in the fully integrated databases, covering more than 245 million patients since 1995 (IBM, 2019). The commercial database primarily consists of employer-sourced and health plan–sourced data containing medical and drug utilization data for several million individuals annually. Medical claims in the commercial database include complete payment and charge information, including amount of patient responsibility. Other standardized items on each medical claim include but are not limited to dates and place of service (e.g., inpatient, outpatient, emergency), diagnoses, procedures, and detailed information on hospitalizations, including admission and discharge dates. Pharmacy claims in the commercial database include complete outpatient prescription drug information, which consists of patient co-payments, mail order drugs, injectables, drugs from specialty pharmacies, and all standardized prescription-level fields collected on a typical pharmacy claim (e.g., date of fill/refill, drug name and class, strength, quantity, and days' supply). The data include co-payment information for inpatient, outpatient, and pharmacy claims. All claims are paid and adjudicated, and the MarketScan research databases fully comply with the Health Insurance Portability and Accountability Act of 1996. Data validation against the original source and/or access to medical records is not available (IBM, 2019).

Maternal and infant records in MarketScan can be linked using the common family identifier, timing of delivery from the mother's file, and birth from the infant's file (MacDonald et al., 2019). The family identifier clusters family members under the same insurance plan. The date of delivery can be derived from the date of delivery procedures in the mother's record, and the date of birth can be approximated by the date of the first claim in the infant's record. In a recent study using MarketScan, this method allowed linkage of 69% of infants (MacDonald et al., 2019). Application of more stringent criteria relating the timing of delivery and birth resulted in the loss of 7% of mother-infant pairs. In the end, 50% of all identified pregnancies (ending in a live birth or otherwise) was linked to an infant.

9.4.4 Danish National Health Databases and the Multiple Sclerosis Registry

The Danish health care system provides universal coverage to all Danish residents (5.6 million inhabitants; http://international.ucl.dk/life-in-denmark/the-danish-health-care-system). Health care coverage includes free visits to general practitioners and specialists, hospital admissions, and outpatient visits. The costs of medications are partially covered by the Danish health system. The centralized Civil Registration System in Denmark allows for personal identification of each person in the entire Danish population and for the possibility of linkage to all Danish registries containing civil registration numbers, such as the Danish National Patient Register, Danish National Prescription Registry, and the Danish Register of Causes of Death. Data collected in these registries are available for research purposes. The process requires collaboration with a local university or investigator affiliated with a research institute to access the data and ethics committee notification or approval to handle data (Danish Data Protection Agency, 2011; Danish Health Authority, 2016). All applications must be submitted in Danish.

Denmark's primary health care sector, which includes general practitioners, specialists, and dentists, generates about 96% of the prescription sales, most of which are reimbursable and are dispensed by community pharmacies. Each dispensing record contains information on the patient, drug, and prescriber. Dispensing records retain the patient's universal personal identifier, allowing for individual-level linkage to all Danish registries and medical databases. Two national registries (Danish National Patient Register and Danish National Prescription Registry) and the Danish National Database of Reimbursed Prescriptions will be of particular interest for implementation of the ocrelizumab post-authorization safety study (PASS). Moreover, the Danish National Civil Registration System will be used to obtain information on death and migration status (Schmidt et al., 2014).

The Danish Medical Birth Registry

This register contains computerized records of all births in Denmark since January 1, 1973. Data are recorded by the midwives or the physicians attending the deliveries

(Knudsen and Olsen, 1998). The registry includes information on maternal age, parity, multiplicity of gestation, birth weight, gestational age, self-reported maternal smoking status, and data about delivery.

The Danish National Patient Register

This register includes data on all hospital admissions since January 1, 1977, and on outpatient clinic and emergency department visits since 1995 (Danish Health Authority, 2016; Lynge et al., 2011; Schmidt et al., 2014). Hospital discharge diagnoses and information on surgical procedures, in-hospital deaths, and some selected drugs are recorded. After 1993, hospital discharge diagnoses are coded using ICD-10 codes.

The Danish National Prescription Registry

Individual-level data on all prescription drugs sold in Danish community pharmacies have been recorded since 1994 in the Register of Medicinal Products Statistics of the Danish Medicines Agency. The register subset, termed the Danish National Prescription Registry, contains information on dispensed prescriptions, including variables at the level of the drug user, the prescriber, and the pharmacy (Kildemoes et al., 2011). The National Prescription Registry collects data on reimbursed and unreimbursed drugs.

The Danish National Database of Reimbursed Prescriptions

This data source encompasses the reimbursement records of all reimbursed drugs sold in community pharmacies and hospital-based outpatient pharmacies in Denmark since 2004 (Johannesdottir et al., 2012). On average, approximately 3.5 million users are recorded in the database each year. Individuals are identified by the unique central personal registration number assigned to all persons born in or immigrating to Denmark. This data source avoids restrictions imposed on data use at the Danish National Prescription Registry. Most importantly, central personal registration numbers are reversibly encrypted, which allows re-identification of drug users. These features are very important for validation purposes.

Danish Multiple Sclerosis Registry

The Danish Multiple Sclerosis Registry (DMSR) was formally established in 1956 and is a nationwide population-based follow-up registry for all Danish patients with MS. Since 1996, the DMSR includes data on all patients who have received DMT. The former Danish Multiple Sclerosis Treatment Registry, established in 1996 to collect data on Danish patients with MS and clinically isolated syndrome who are treated with DMTs, is now integrated into the DMSR. Reporting of patients on DMTs is mandatory (Flachenecker et al., 2014; Koch-Henriksen et al., 2015; Koch-Henriksen et al., 2012; Magyari et al., 2016). By regular data collection, the DMSR makes it possible to follow the therapeutic effect of DMTs such as relapses, adverse effects, and disability (Bronnum-Hansen et al., 2011). At present, more than 11,000 patients treated with DMT have been registered in this. database. Data are continuously entered into a central database from all sites in Denmark at start and at regular visits. Population-based

studies combining health and social registries can optimally be carried out in Denmark, due to linkage between registries at the individual level by a unique personal identification number (CPR number), which is used by all national registries. Thereby, the DMSR can be linked to a number of population-based registries to obtain additional patient information, such as hospitalization, cause of death, comorbidities, prescription drugs, and reproductive issues. The minimum data set of DMSR includes information about age, sex, year of disease onset, year of first diagnosis, basic clinical information, and information about treatment, side effects, relapses, and patient neurostatus expressed by Functional Systems Score (FSS) and Expanded Disability Status Scale (EDSS) score. Notification is done at treatment start, and thereafter at every scheduled clinical visit, 3 months after treatment start and thereafter every 6 months. The longitudinally collected information about disease activity and side effects make it possible to investigate the clinical efficacy and adverse events of different DMTs.

Due to the unique CPR number in Denmark, patient follow-up is lifelong, as long as the patient remains in the country. Currently, serious adverse events are available through linkage to other registries. Prospective inclusion of adverse drug reaction or serious adverse events as part of the minimum data set will be possible, due to an expansion, via addition of a safety module, of the online data collection platform COMPOS DK.

9.5 STUDY SIZE

The target size for this study was calculated so that the upper limit of the 95% confidence interval for the hazard ratio for the association of pregnancy exposure to ocrelizumab and major congenital malformations is below 2.5 with 0.8 probability. This number is estimated at approximately 1,005 exposed pregnancies and 3,015 unexposed pregnancies in each of the two comparator cohorts, counting subjects from the four proposed data sources. Figures are as follows:

- 1,005 exposed and 3,015 unexposed pregnancies (in each of the two comparator cohorts) would be accrued combining the four data sources
- Of these, 62% would result in live births (623 exposed and 1,869 unexposed)
- Of these records, 65% would be linkable to infant records, resulting in 405 exposed and 1,215 unexposed newborns, per Table 5

Actual study size will depend on medication utilization among the target population in the data sources selected for the study, as well as on the observed percentage of live births among the study pregnancies and success of mother-infant record linkage.

The estimations for ocrelizumab exposure, potentially the limiting factor for reaching the study target size in this study, are described in Table 5 and associated text, below.

Table 5 Study Sizes Needed to Have Probability of 0.8 That the Upper Limit of the 95% Confidence Intervals Will Be Below Selected Thresholds

Outcome	Prevalence of Outcome	Upper Limit of 95% CI for Relative Risk Will Be Less Than	Exposed:Unexposed Pregnancies Needed
Stillbirth	6 per 1,000[a]	11	300:900
		8	400:1,200
		5	675:2,025
		3	1,440:4,320
Cardiac congenital malformations	1%[b]	12.75	160:480
		8.99	215:645
		6.5	300:900
		5	400:1,200
		4	540:1,620
Major congenital malformations (combined)	3%[c]	5	130:390
		4.3	160:480
		4	177:531
		3.5	215:645
		3	280:840
		2.9	300:900
		2.5	405:1,215
		2	705:2,115
Preterm birth	10%[d]	5	37:111
		4	50:150
		3	80:240
		2.2	160:480
		1.95	215:645
		1.85	250:750
		1.75	300:900
		1.5	575:1,725

CDC = Centers for Disease Control and Prevention; CI = confidence interval; MS = multiple sclerosis.

Note: Assumptions underlying these calculations:

- No difference in risk between the exposed and unexposed (i.e., risk ratio = 1), regardless of comparison cohort (women with MS without ocrelizumab exposure, women without MS).
- Matching ratio of exposed to unexposed was 1:3.
- Probability that the upper limit of 95% CI will be as stated = 0.8.
- Calculations were done using the "Study Size" tool in Episheet (Rothman, 2015).

[a] MacDorman and Gregory (2015).

[b] CDC (2016a).

[c] CDC (2008).

[d] CDC (2016c).

The potential number of pregnancies exposed to ocrelizumab per year was estimated using counts provided by the database holders, several assumptions related to pregnancy rates in the general population and in women with MS, and the projected average use of ocrelizumab (Table 6). The potential accrual of pregnancies per year presented in Table 6 represents approximate estimates with many uncertainties. Concretely, the pregnancy rate may be lower than assumed (based on the general population), and the rate of treatment discontinuation before pregnancy can be higher than assumed (50%).

Each year, the observed mean number of women with MS with a delivery code in Optum DAPI and who used DMTs currently approved for the treatment of MS for years 2012-2016 were consistent with estimates for ocrelizumab in Table 6 (data provided by Optum).

Table 6 Estimated Number of Pregnancies Exposed to Ocrelizumab in Women With Multiple Sclerosis in Selected US Data Sources, per Year

Data Source	Number of Women Aged 15-45 Years With MS per Year	Estimated Number of Pregnan-cies per Year[a]	Projected Number of Pregnancies Exposed to Ocrelizumab per Year[b]	Adjusted to 50% Withdrawn From Ocrelizumab Before Pregnancy[c] (per year)	Adjusted Assuming 10% Loss to Follow-up (per Year)	Adjusted Assuming 30% Loss to Follow-up (per Year)
HIRD (US)	31,295 (Jan 2006-Apr 2016) ~9,400 per year	990	57	29	26	20
Market-Scan (US)	25,729 (2014)	2,700	157	79	71	55
DAPI, Optum (US)	7,421 (2015)	780	45	23	21	16
Danish Registries		73 (observed average 2007-2012)	4	2	2	1
Total	--	4,543	263	133	120	92

HIRD = HealthCore Integrated Research Database℠; MS = multiple sclerosis; DAPI = Optum Dynamic Assessment of Pregnancies and Infants™; US = United States.

[a] Assuming a pregnancy rate of 10.5% in US women aged 15-44 years (Ventura et al., 2012).

[b] Assuming an average proportion of patients with MS treated with ocrelizumab of 5.8%, which could be slightly lower if use is strictly restricted to the approved indications of primary progressive and relapsing forms of MS.

[c] Assuming a 50% lower estimated number of pregnancies exposed due to the possible decision to stop treatment before becoming pregnant.

Individually, none of these data sources would provide a sufficient number of pregnancies with potential exposure to ocrelizumab by 5 years after launch to provide adequate power to draw meaningful conclusions. Hence the multidatabase approach and the need to monitor the actual number of pregnant women in each database during the first years to confirm the feasibility and timing of the study start.

Per the estimated 92 exposed pregnancies that may be observed yearly (Table 6), attaining the estimated target study size would take approximately 11 years. All eligible pregnancies observed during the study period will be included in the study. Because some of the assumptions used in these calculations may be too conservative (for example, in Danish data, practically all deliveries are linkable to infant records), it is possible that this number of pregnancies is observed earlier, and analyses can proceed earlier than planned.

A common monitoring plan and a common SAP will guide analyses. The monitoring plan will be developed before data for the first monitoring report are extracted, will describe the analyses to be conducted periodically for monitoring accrual of ocrelizumab-exposed pregnancies and live births, and will update the predicted study power. Briefly, counts of ocrelizumab-exposed pregnancies and live births will be obtained from each data source using the most recently updated data available; the live birth proportion in each data source will be estimated. The SAP will be developed before database lock and data extraction and will provide details on the study core analyses and meta-analysis. Data from monitoring reports will inform the SAP.

The number of ocrelizumab-exposed pregnancies and live births will be monitored yearly in all databases in years 2020 through 2028. Results from monitoring counts, along with any updated figures on the success rate of mother-infant linkage in the US data sources, will allow study researchers to estimate when the target study size will be attained and the anticipated duration of the study.

FDA has recommended that the study be extended as needed to attain the target study size. If the target study size is achieved, but the number of pregnancies available in one data source is not large enough to support the planned analyses, descriptive or unadjusted analyses will be considered for that data source.

9.6 DATA MANAGEMENT

Files from the various data sources will be kept separate behind firewalls, with each research partner analyzing data from different data sources, and individual-level data will not be merged.

Routine procedures will include checking electronic files, maintaining security and data confidentiality, following analysis plans, and performing quality-control checks of all programs. The study is conducted by multiple research partners, and each research partner will maintain any patient-identifying information securely on site according to

internal/local standard operating procedures or guidance documents. Security processes will be in place to ensure the safety of all systems and data. Every effort will be made to ensure that data are kept secure so that they cannot be accessed by anyone except selected study staff.

Appropriate data storage and archiving procedures will be followed, with periodic backup of files. Standard procedures will be in place at each research center to restore files in the event of a hardware or software failure.

Each research partner will follow its own established procedures and generate results according to the analysis plan and specifications. All summary tables of results, and no individual patient identifiers, will be provided to RTI Health Solutions (RTI-HS), which will compile the results and develop the report in collaboration with research partners. RTI-HS will follow quality-control procedures regarding transfer of data.

For requests to access to data for audit purposes, only aggregated data from all research centers will be available at the coordinating center. The audit trail will consist of a detailed description of the methods to extract and process the electronic health records or claims data, as applicable. Access to raw data at each database research center will require the data requestor to obtain a license or apply for approval at a research committee and to fulfill the conditions required under the governance rules of each database research center.

9.7 DATA ANALYSIS

9.7.1 Study Cohorts

The study cohorts are summarized in Table 7.

Table 7 Study Cohorts

	Exposed Cohort	Primary Comparison Cohort		Secondary Comparison Cohort
		Comparison Subcohort 1a	Comparison Subcohort 1b	
Cohort definition	Pregnancies in women with MS and exposure to ocrelizumab	Pregnancies in women with MS exposed to non-ocrelizumab DMTs approved for the treatment of MS	Pregnancies in women with MS not exposed to DMTs approved for the treatment of MS	Pregnancies in women without MS or ocrelizumab use
MS diagnosis	Required	Required	Required	Absence of diagnosis required
Drugs	Required: Use of ocrelizumab in the 6 months before pregnancy or any time during pregnancy (first trimester only for analysis on MCM)	Required: Use of DMTs in the 6 months before pregnancy or any time during pregnancy (first trimester only for analysis on MCM) No use of ocrelizumab in the 6 months before pregnancy or any time during pregnancy (first trimester only for analysis on MCM)	No use of DMTs in the 6 months before pregnancy or any time during pregnancy (first trimester only for analysis on MCM)	No use of ocrelizumab in the 6 months before or any time during pregnancy

DMT = disease-modifying therapy; MCM = major congenital malformation; MS = multiple sclerosis.

In Denmark, MS will be ascertained from the MS Registry. The ascertainment of MS in claims is not straightforward. For US data, based on algorithms in the literature (Table 8), it is proposed to consider women to have MS if they have at least one claim for MS and one claim for an MS DMT, an MS-related sign or symptom (e.g., optic neuritis), or an MS-related service in their baseline period or during pregnancy. Codes during pregnancy are included because they are understood to represent a diagnosis that was present before pregnancy, but is recorded only during pregnancy. To define absence of MS (secondary comparison cohort), it will be required that patient records do not contain MS diagnosis codes or entries for DMTs to treat MS.

Table 8 Published Algorithms to Identify Cases of Multiple Sclerosis in Claims Data

Reference, Database, and Country	Summary of Algorithm	Validation Results
Bargagli et al. (2016) Health administrative data including the Hospital Discharge Registry and the PHARMED database of dispensed medications within the Italian National Health System; Italy	Selected patients with hospital records with a primary or a secondary diagnosis of MS (ICD-9-CM, 340.0); with at least one pharmacy claim in the study period for at least one MS DMT; and patients with records of MS-related disability.	No validation results.
Culpepper et al. (2006) Several data sets, including administrative and health data and electronic medical charts, from the Veterans Health Administration data sources; United States	The database algorithm classified patients as not having MS if they did not have, on average, at least one health care encounter each year with MS coded as the primary diagnosis (ICD-9-CM, 340), did not have a service-connected disability for MS (this was related to compensation and pensions), or did not use a DMT. If any of these criteria were met, the case was classified as MS/possible MS.	In validation against medical charts, sensitivity was 0.93; specificity, 0.92; positive predictive value, 0.92; and negative predictive value, 0.93. In an analysis in which cases classified as unknown in the medical chart review were retained, sensitivity was 0.93; specificity, 0.90; positive predictive value, 0.88; and negative predictive value, 0.94.
Higuera et al. (2016) Claims data from commercial plans, individual and family health plans, managed Medicaid and Medicare plans, and dual Medicaid/Medicare plans in the upper Midwest of the United States.	Patients had to have an ICD-9 code for MS (340 or 340.0) and a prescription filled for an MS DMT in any year of data (this was a study on adherence to DMTs)	No validation results.

Reference, Database, and Country	Summary of Algorithm	Validation Results
Marrie et al. (2010) Health claims from the Manitoba Health and Healthy Living program—including physician claims, hospitalization records and outpatient dispensed prescriptions—from the Drug Programs Information Network; Canada.	The authors validated several algorithms requiring different numbers of and combinations of claims for hospitalization, general practitioner encounters, and prescriptions. The base cohort from which cases were selected had codes for signs and symptoms that may be related to MS. One version of the algorithm required at least two MS claims (physician, hospital, or DMT prescriptions). ICD-9-CM and ICD-10-CA codes used for MS were 340 and G35, respectively. A simpler version required one claim of any of the three types.	Validation was conducted against medical records. Two claims required: sensitivity was 95.7%, sensitivity, 44.3%; PPV, 73.2%; and NPV, 86.7%. One claim required: sensitivity was 97.9%, specificity, 17.0%; PPV, 65.2%; and NPV, 83.3%.

DMT = disease-modifying therapy; ICD-10-CA = *International Classification of Diseases, 10th Revision, Canadian Modification*; ICD-9 = *International Classification of Diseases, 9th Revision*; ICD-9-CM = *International Classification of Diseases, 10th Revision, Clinical Modification*; MS = multiple sclerosis; NPV = negative predictive value; PPV = positive predictive value.

9.7.2 Descriptive Analyses

In each data source, the study cohorts (exposed cohort, comparison cohort 1, subcohorts 1a and 1b, and comparison cohort 2) will be characterized based on the covariates listed in Section 9.3.3. Characteristics of the unmatched and matched cohorts, and frequency of the endpoints, will be shown in a table. Balance in matching will be assessed by examining the distribution of variables in the cohorts and the standardized difference. No statistical tests are planned for this comparison, but variables with standardized differences above 0.1 may lead to a re-evaluation of the propensity score estimation.

9.7.3 Measures of Frequency

In each data source, crude measures of incidence or prevalence of the study outcomes with associated 95% confidence intervals will be estimated within matched cohorts (exposed cohort, primary comparison cohort, subcohorts 1a and 1b, and secondary comparison cohort; see next subsection).

9.7.4 Measures of Association

Table 9 presents the measures of association planned for each outcome. Regression models will be used to compare women with MS who received ocrelizumab during the exposure window with women in the primary comparison cohort and subcohorts 1a and 1b, and in the secondary comparison cohort, which have been described above. Point estimates and 95% confidence intervals from crude analyses within the matched cohorts will be presented. The number of pregnancies will be considered when determining which analyses can be conducted.

To control for confounding and channeling effect (see Section 9.9), women in the ocrelizumab-exposed cohort will be matched to women in the primary and secondary comparison cohorts (separately) in a 1:3 ratio in a variable-matching ratio based on an exposure propensity score (i.e., if only one pregnancy from the primary comparison cohort is available for a given exposed pregnancy, that matched set will have one exposed pregnancy and one comparison pregnancy), using greedy matching.

An advantage of matching is that “crude” results are adjusted for the matching variables. The main disadvantage is the loss of precision associated with the loss of unmatched subjects in the context of a rare exposure and rare outcomes. However, because many more pregnancies are expected in the two comparison cohorts unexposed to ocrelizumab, it is unlikely that ocrelizumab-exposed pregnancies will be left without appropriate matches. Therefore, this disadvantage is not a major concern in this setting. The variable-matching ratio has the advantage of minimizing the loss of exposed pregnancies due to lack of matches while increasing precision due to multiple matches for easily matchable exposed pregnancies.

Logistic regression will be used to estimate an exposure propensity score as the probability of being exposed to ocrelizumab or not (separately for the primary and secondary comparison cohorts). Variables considered for inclusion in the propensity score will be those related to the exposure and to any of the outcomes; the method for variable selection will be specified in the SAP. The propensity score model for the secondary comparison cohort will be a reduced model to avoid selecting a cohort composed of atypical pregnancies from the general population that have many of the characteristics of ocrelizumab-exposed pregnancies with MS but do not have exposure to ocrelizumab or a diagnosis of MS. Details will be provided in the SAP.

A strength of propensity scores is that they allow for the selection of the best set of variables within each data source. For example, Danish data will be richer with regard to information on MS (to be used in propensity scores in cohorts of women with MS only [exposed cohort and primary comparison cohort]). Propensity scores will allow the investigators to use this information to obtain a locally optimal control of confounding. Furthermore, flexibility in variable selection may allow the investigators to reduce overall confounding in this setting in which confounding factors will likely vary across data sources, given differences in insurance, health care systems, and patient characteristics. Propensity scores with subsequent pooling of data source–specific information have been used in multidatabase studies (Rassen et al., 2010; Toh et al., 2013). A single propensity score will be used for all outcomes in comparisons of ocrelizumab-exposed pregnancies and pregnancies in the primary comparison cohort (then, for comparison subcohorts 1a and 1b), and another propensity score will be used for all outcomes in comparisons of ocrelizumab-exposed pregnancies and pregnancies in the secondary comparison cohort.

9.7.5 Missing Data

Using automated health care data, missing data for exposure, outcome, and comorbidities are not expected; in the presence of records for the condition, it is assumed the condition is present, and in the absence of such records, it is assumed that the condition is absent. Otherwise, where relevant, the percentage of missing data will be reported.

9.7.6 Statistical Analyses

A summary of the analyses proposed for each outcome can be found in Table 9. The feasibility of the proposed approaches will be reassessed taking into account the number of observed outcomes.

For analyses of urinary tract infections in pregnancy and infections requiring hospitalization during pregnancy, person-time will accrue from the beginning of the pregnancy until the end of the woman's follow-up, as described in Section 9.2.3, Follow-up. For analyses on adverse effects on the infant immune system, i.e., vaccine-preventable diseases and vaccine-associated poliomyelitis in the infant, person-time will

accrue from the date of birth until the end of the infant's follow-up, as described in Section 9.2.3, Follow-up, and in Table 9, in the column "Timing of Outcome Ascertainment" (this column notes that some person-time analyses will have multiple cut-off ages). In person-time analyses, exposure will be allowed to vary over time from unexposed to exposed. Once exposed, person-time will continue to be considered exposed because the potentially deleterious effects of drugs may persist after the drug has been cleared from the body (e.g., the depletion of CD20+ B cells may persist after a CD20+ B-cell–depleting drug has been catabolized or excreted).

Table 9 Proposed Statistical Analyses

Outcome	Measure of Frequency	Measure of Association (Regression Model)	Timing of Outcome Ascertainment	Timing of Exposure Ascertainment	Unit of Analysis
Spontaneous abortion	Cumulative risk	Odds ratio (logistic regression)	Before week 20 of pregnancy	6 months before conception up to and including first trimester of pregnancy, or end of pregnancy if prior to end of first trimester	Pregnancy
Fetal death/ stillbirth	Cumulative risk	Odds ratio (logistic regression)	At or after week 20 of pregnancy	6 months before conception or any time in pregnancy (i.e., before the time of fetal death)	Pregnancy
Elective termination	Cumulative risk	Odds ratio (logistic regression)	At any time of follow-up	6 months before conception up to and including first trimester of pregnancy, or end of pregnancy if prior to end of first trimester	Pregnancy
Preterm delivery	Cumulative risk	Odds ratio (logistic regression)	At birth	6 months before conception or any time in pregnancy	Pregnancy
C-section	Cumulative risk	Odds ratio (logistic regression)	At birth	6 months before conception or any time in pregnancy	Pregnancy
Urinary tract infection	Incidence rate	Hazard ratio (proportional hazards regression)	At any time in pregnancy (or end of follow-up for the patient)	6 months before conception until the end of follow-up for the patient for this outcome	Pregnancy
Infection requiring hospitalization during pregnancy	Incidence rate	Hazard ratio (proportional hazards regression)	At any time in pregnancy (or end of follow-up for the patient)	6 months before conception until the end of follow-up for the patient for this outcome	Pregnancy

Outcome	Measure of Frequency	Measure of Association (Regression Model)	Timing of Outcome Ascertainment	Timing of Exposure Ascertainment	Unit of Analysis
Major congenital malformations	Cumulative risk	Odds ratio (logistic regression)	At birth or during infant follow-up	6 months before conception up to and including first trimester of pregnancy	Fetuses (live births or others if the information is available for terminations or stillbirths) or infants
Adverse effects on the immune system	Incidence rate	Hazard ratio (proportional hazards regression)	The hazard ratio will be evaluated at several ages, because the incidence may vary widely with age (e.g., first month of life, and at the end of infant follow-up). Subgroup analyses on infants who died because of infectious disease or had a full year of follow-up may be explored. For vaccine-preventable diseases and vaccine-associated poliomyelitis, the outcome will be ascertained between age for first dose of vaccines and end of first year of life.	6 months before conception or any time in pregnancy	Infants
Vaccine-preventable diseases and vaccine-associated poliomyelitis	Incidence rate	Hazard ratio (proportional hazards regression)	The hazard ratio will be evaluated at several ages because the incidence may vary widely with age (e.g., first month of life, and at the end of infant follow-up). The outcome will be ascertained between age for first dose of vaccines per guidelines and end of first year of life.	6 months before conception or any time in pregnancy	Infants

Outcome	Measure of Frequency	Measure of Association (Regression Model)	Timing of Outcome Ascertainment	Timing of Exposure Ascertainment	Unit of Analysis
Small for gestational age	Prevalence	Odds ratio (logistic regression)	At birth	6 months before conception or any time in pregnancy	Fetuses (live births)

Note: With variable-matching ratios, analyses such as logistic regression would be conditional on the matched sets. Pregnancies no longer at risk for one outcome will be excluded (e.g., a pregnancy that ended in a spontaneous abortion is not at risk for preterm delivery). The date or gestational age at the time of the event may need to be estimated for spontaneous abortions, elective terminations, and stillbirth, as generally they are not well recorded. For each outcome of interest that can occur multiple times, follow-up for that outcome will stop at its first occurrence (e.g., urinary tract infections in pregnancy, infections requiring hospitalization in pregnancy). Follow-up will continue for other outcomes.

9.7.7 Data Integration

Results will be presented separately for each data source. Data analysis will be performed by data custodians at their sites and behind firewalls, and individual-level data will not be available for data integration. Overall results (e.g., odds ratios for MCMs) will be summarized using meta-analytic techniques. It is proposed to use direct meta-analytic techniques (meta-analytic techniques with direct comparisons) with random effects as results are expected to have some heterogeneity across data sources; the random-effects model will allow assessment of heterogeneity. Standard meta-analysis graphic displays and diagnostics (Forest plots and I^2) will be conducted to assess heterogeneity of effect across data sources.

Outcomes for data integration include all outcomes listed, plus subgroups of MCMs (e.g., cardiac malformations, cleft lip with or without cleft palate). Results from sensitivity analyses (e.g., using various exposure windows) may be pooled if numbers are adequate.

9.7.8 Subgroup Analyses

Possible stratifications (depending on counts) may include strata of maternal age (e.g., 20-24 years, 25-29 years), calendar year, and others.

9.7.9 Sensitivity Analyses

Ocrelizumab has a mean half-life of 26 days. It is expected that after 5 half-lives (i.e., 130 days) after the last ocrelizumab administration, ocrelizumab would be eliminated from the women's bodies. Accordingly, the following sensitivity analyses related to the definition of unexposed are proposed:

- Women would be considered unexposed if their last recorded exposure to ocrelizumab took place 26 or more days before the beginning of pregnancy (1 mean half-life).
- Women would be considered unexposed if their last recorded exposure to ocrelizumab took place 90 or more days before the beginning of pregnancy (3.5 mean half-lives).
- Women would be considered unexposed if their last recorded exposure to ocrelizumab took place 130 or more days before the beginning of pregnancy (5 mean half-lives).

Other sensitivity analyses will consider last exposure of ocrelizumab took place within the first, second or third trimester of pregnancy. Exposure during the second and third trimester may be combined for the outcome major malformation to maximize study size, taking into consideration that the critical time window for this outcome is the first trimester.

Other sensitivity analyses will be planned around the handling of covariates.

9.8 QUALITY CONTROL

Data management and analysis will be conducted by each research partner in each database. Standard operating procedures or internal process guidance at each research center will be used to guide the conduct of the study. These procedures include internal quality audits and the opportunity for external audits; rules for secure and confidential data storage, backup, and recovery; methods to maintain and archive project documents; quality-control procedures for programming; standards for writing analysis plans; and requirements for scientific review by senior staff. A quality-assurance audit of this study may be conducted by the sponsor or the sponsor's designees. Each of the database research centers will follow its own quality and audit trail procedures. The quality and audit trails at each center may be different.

At the coordinating center, an independent Office of Quality Assurance performs audits and assessments that involve various aspects of its projects, including but not limited to documentation of education and training, data entry, data transfer, and approval of the institutional review board at RTI International, of which RTI-HS is a research unit. Such audits at RTI-HS will be conducted by the Office of Quality Assurance according to established criteria in standard operating procedures and other applicable procedures.

All programming written by one study analyst will be reviewed independently by a different analyst, with oversight by a senior statistician. All key study documents, such as the analysis plan, data abstraction forms, and study reports, will undergo quality-control review, senior scientific review, and editorial review.

External reviewers with expertise in drug safety in pregnancy and MS provided advice on the design of the study and have reviewed this protocol; they will provide advice on the conduct of the study and will review results, reports, and other important documents.

9.9 LIMITATIONS OF THE RESEARCH METHODS

Data-related limitations of the study will depend on the accuracy of codes and algorithms used to identify outcomes, completeness of the data contained in the data sources, and the availability of records for validation of selected outcomes. Of the three health care claims databases included in this study, only MarketScan does not have access to medical records. Optum has the ability to access medical records for data abstraction and validation of study endpoints, and claims data in HIRD can be linked to medical records. Definitions for and algorithms to identify study endpoints will be homogeneous across databases; given that the same algorithm will be applied and that US claims data sources share basic features, positive predictive values obtained from DAPI and HIRD are expected to be applicable to MarketScan. In Denmark, it is possible to apply for access to the medical files to retrieve additional data or to validate registry-based information, if needed.

Not all of the outcomes of interest have been validated in administrative claims data, and the performance of ICD-10 codes, which have only been used since October 2015 in the US, has not been well characterized in this setting.

Exposure ascertainment will be based on pharmacy claims in US data sources. Given the type of medication, it is expected that filled prescriptions will be used (as opposed to use-on-demand medications, such as pain killers), although the precise timing of use will be uncertain. In Denmark, exposure ascertainment will be more robust thanks to the MS register. Exact timing of exposure relative to start of pregnancy will need to be estimated in claims because claims data sources do not include date of last menstrual period or records of gestational age at birth. However, multiple methods for this estimation using data available in claims have been evaluated and found to be appropriate (Margulis et al., 2015). Information on breastfeeding will be limited or unavailable.

Unsuccessful linkage of mothers and infants may reveal some differences between mothers linked to infants and mothers without linked infants. Bias due to non-linkage is not expected unless the characteristics associated with linkage or lack thereof are simultaneously associated with the exposure and the outcomes, which is not likely.

Similar to the design other pregnancy safety studies (e.g., pregnancy registries), early pregnancy outcomes, such as spontaneous abortions, will be incompletely captured in pregnancy studies based on existing databases, because they may happen before the pregnancy is known to the health care system. Information related to pregnancy terminations and reasons for termination will be incomplete. Outcomes that occur during the first year of life will be captured only in infants who stay enrolled for that time period.

Information on some potential confounding factors, such as use of supplements, tobacco, alcohol and illicit drugs, is limited in these data sources. Information on lifestyle factors such as smoking, alcohol consumption, and body mass index is not typically available in claims databases. However, the presence of diagnoses related to obesity, smoking, and alcohol abuse or dependence can be used to approximate the patient profile relating to health habits in databases where this information is missing. Danish data include some lifestyle information, as it is captured and recorded during prenatal care. In claims data, the period for ascertaining baseline characteristics will be limited by the period of enrollment in the health care plan; underrecording of diagnoses only recorded well before pregnancy is expected. For certain diagnosis that are mainly recorded at the primary care level, some degree of underascertainment is expected in the Danish databases if the patients did not receive inpatient or outpatient hospital care for those conditions. Over-the-counter medications, including folic supplementation at the regular dose and prenatal vitamin use, are not captured in these data sources. Claims data will also provide limited or no information on type of MS, duration of the disease, and other MS-related details. Some degree of residual confounding due to this can be expected, although the proposed method to adjust for confounding should remove most confounding.

Currently, several drugs are available to treat relapsing MS (ocrelizumab is the first medication approved for the indication of PPMS). At this point, one can only speculate about which patients would receive prescriptions for ocrelizumab. Matching on patient characteristics will aim to minimize confounding due to channeling effect, although it will not remove it completely. Because of the Danish MS Registry, Danish data will be more robust in this respect. This study is not designed to study the effect of ocrelizumab in patients without MS or patients with isolated clinical syndromes who do not have a diagnosis of MS.

Due to the low frequency of exposure and outcomes, this study will reach limited precision within a reasonable time frame. While a smaller study size (e.g., a minimum of 100 ocrelizumab-exposed pregnancies) would ensure the study is conducted even in the case of low uptake of ocrelizumab, results would be very imprecise for all but the most common study outcomes. Subcohorts 1a and 1b, while offering comparisons of clinical interest, will also be limited by low counts of participants and outcomes.

It is possible that some pregnancies are represented in more than one participating US claims databases. Membership in the HIRD and DAPI is health care plan–based, while membership in MarketScan is employer-based. Thus, women and infants can be represented in either HIRD or DAPI and also in MarketScan if health care is provided by an employer represented in MarketScan and through a health insurance plan represented in HIRD or DAPI. No information is directly available in any of the data sources to identify the overlap in data sources.

It is possible that women with MS and their children are subjected to more intensive surveillance during pregnancy and in childhood, respectively, possibly resulting in higher recorded prevalence of non–life-threatening malformations than women without MS. If this is true, we should expect to see higher prevalence of malformations in infants born to women with MS, even if MS or MS treatments do not carry such increased risk. If increased surveillance is not differential by MS treatment, comparisons between ocrelizumab and other DMTs approved for MS should not be affected.

10. PROTECTION OF HUMAN SUBJECTS

This is a non-interventional study using secondary data and does not pose any risks for patients. All data collected in the study will be de-identified with no breach of confidentiality with regard to personal identifiers or health information. Each US database research partner will apply for an independent ethics committee review according to local regulations; in addition, as the coordinating center, RTI-HS will obtain approval from the RTI International institutional review board (IRB) (see Section 10.1). In Denmark, following local regulations, approval from the Data Protection Authorities will be obtained. Data protection and privacy regulations will be observed in collecting, forwarding, processing, and storing data from study participants.

In the case of secondary data collection for validation purposes (e.g., medical chart abstraction), every effort will be made to protect participant confidentiality according to local and international data privacy and medical record confidentiality guidance and requirements.

10.1 RTI INTERNATIONAL

RTI International holds a Federal-wide Assurance from the Department of Health and Human Services Office for Human Research Protections that allows the organization to review and approve human subjects protocols through its IRB committees. RTI International currently has two IRB committees available to review research protocols. One IRB committee is constituted to review medical research and has two members who are physicians. These IRBs have been audited by the FDA and are fully compliant with applicable regulatory requirements.

10.2 HEALTHCORE INTEGRATED RESEARCH DATABASE℠

HealthCore maintains Data Sharing Agreements and Business Associate Agreements with all covered entities who provide data to the HIRD. HealthCore's access, use, and disclosure of protected health information (PHI) are in compliance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule (45 CFR Part 160 and Subparts A and E of Part 164). HealthCore does not access, use, or disclose identifiable PHI unless under a specific waiver of authorization (e.g., a HIPAA Waiver of Authorization from an IRB). HealthCore accesses the data in a manner that complies with federal and state laws and regulations, including those related to the privacy and security of individually identifiable health information.

As PHI must be accessed to acquire medical records to validate electronic case-finding algorithms, a HIPAA Waiver of Authorization will be applied for from an IRB. HealthCore will submit the protocol to a central IRB for review and approval. If changes to the protocol are required, HealthCore will submit an amendment to the IRB.

At no time during the conduct of this study will HealthCore provide patient- or provider-identifying information to RTI-HS or Roche. Only aggregated data will be reported to RTI-HS or Roche. HealthCore will keep all needed documents and provide a copy of the final approvals/waivers to RTI-HS (the coordinating center) for completeness of study records.

10.3 OPTUM DYNAMIC ASSESSMENT OF PREGNANCIES AND INFANTS™

Optum will apply for approval to health plans and an IRB as per standard procedures. Optum will communicate directly with the IRB to address questions or provide additional information as needed. The IRB will be asked to re-approve the study with the periodicity established by standard procedures. In addition to IRB approval, internal review and approval are also required. Optum will keep all needed documents and provide a copy of

the final approvals/waivers to RTI-HS (the coordinating center) for completeness of study records.

10.4 IBM MARKETSCAN COMMERCIAL CLAIMS AND ENCOUNTERS DATABASE

RTI-HS will obtain approval from the RTI International IRB.

10.5 DANISH NATIONAL HEALTH DATABASES AND DANISH MULTIPLE SCLEROSIS REGISTRY

For the Danish national health databases, approval will be requested from the Danish Data Protection Agency.

10.6 OTHER GOOD RESEARCH PRACTICE

This study adheres to the *Guidelines for Good Pharmacoepidemiology Practices (GPP)* of the International Society for Pharmacoepidemiology (ISPE, 2015) and has been designed in line with the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) *Guide on Methodological Standards in Pharmacoepidemiology* (ENCePP, 2018b). The ENCePP Checklist for Study Protocols (ENCePP, 2018a) is included in Appendix 2.

The study is a PASS and will comply with the definition of the non-interventional observational study referred to in the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use tripartite guideline Pharmacovigilance Planning E2E (ICH, 2004) and provided in the EMA *Guideline on Good Pharmacovigilance Practices (GVP) Module VIII: Post-Authorisation Safety Studies* (EMA, 2017b), and with the 2012 European Union pharmacovigilance legislation, adopted June 19, 2012 (European Commission, 2012). The study will comply with the study reporting requirements specified in Module VIII Section VIII.B.6.3.2. “Final Study Report” of the GVP (EMA, 2017b).

The study will be registered in the EU PAS Register (ENCePP, 2016) before the study implementation commences.

11. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

For non-interventional studies that are solely based on secondary use of data, reporting of adverse events/adverse drug reactions is not required. Based on current guidelines from the ISPE (ISPE, 2015) and the EMA GVP Module VI - Management and Reporting of Adverse Reactions to Medicinal Products (EMA, 2017a), non-interventional studies such as the one described in this protocol conducted using aggregated patient data from electronic health care records do not require expedited reporting of suspected adverse events/reactions. Based on the data planned for this study, no suspected adverse events/reactions are expected.

12. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

Regardless of the outcome of this study, Roche is dedicated to openly providing information on this safety study to health care professionals and to the public, both at scientific congresses and in peer-reviewed journals. Roche will comply with all requirements for publication of study results. Roche will submit all study reports to the health authorities through scheduled regulatory safety reporting (periodic adverse drug experience report [PADERs] and/or periodic benefit-risk evaluation report [PBRERs], as agreed with health authorities).

The study protocol and progress and final study reports will be included in regulatory communications in line with the risk management plan, PADERs/PBRERs, and other regulatory milestones and requirements. Study reports will be prepared using a template following GVP, Module VIII, Section B.6.3 (EMA, 2017b).

Any publications will follow guidelines, including those for authorship, established by the International Committee of Medical Journal Editors (ICMJE, 2018). When reporting results of this study, the appropriate STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist will be followed (STROBE statement, 2007).

In the contracts for the implementation of this study, Roche and the principal investigators (e.g., the principal investigators at the study coordinating center and at the research partner centers) will agree upon a publication policy allowing the principal investigators to independently prepare publications based on the study results, irrespective of data ownership and sponsorship.

13. REFERENCES

ACOG. Management of stillbirth. American College of Obstetricians and Gynecologists; March 2009.

Anthony M, Krueger W. Pregnancy studies in drug-exposed populations with multiple sclerosis: a literature review. Final report prepared by RTI Health Solutions for F. Hoffmann-LaRoche Ltd. 2016.

Anthony M, Krueger W, Saltus K. Suitable databases for the study of pregnancy and infant outcomes in ocrelizumab-exposed women with multiple sclerosis. Final report prepared by RTI Health Solutions for F. Hoffmann-LaRoche Ltd. 2016.

Bargagli AM, Colais P, Agabiti N, Mayer F, Buttari F, Centonze D, et al. Prevalence of multiple sclerosis in the Lazio region, Italy: use of an algorithm based on health information systems. J Neurol. 2016 Apr;263(4):751-9.

Broms G, Granath F, Linder M, Stephansson O, Elmberg M, Kieler H. Birth outcomes in women with inflammatory bowel disease: effects of disease activity and drug exposure. Inflamm Bowel Dis. 2014 Jun;20(6):1091-8.

Bronnum-Hansen H, Koch-Henriksen N, Stenager E. The Danish Multiple Sclerosis Registry. Scand J Public Health. 2011 Jul;39(7 Suppl):62-4.

CDC. Congenital heart defects (CHDs). Centers for Disease Control and Prevention; 2016a. Available at: https://www.cdc.gov/ncbddd/heartdefects/data.html. Accessed July 13, 2017.

CDC. National Death Index (factsheet). Atlanta, GA: National Center for Health Statistics, Centers for Disease Control and Prevention; 2016b. Available at: http://www.cdc.gov/nchs/data/factsheets/factsheet_ndi.htm. Accessed July 13, 2017.

CDC. Preterm birth. Atlanta, GA: Centers for Disease Control and Prevention; 2016c. Available at: https://www.cdc.gov/reproductivehealth/maternalinfanthealth/pretermbirth.htm. Accessed July 13, 2017.

CDC. Update on overall prevalence of major birth defects—Atlanta, Georgia, 1978-2005. MMWR Morb Mortal Wkly Rep. 2008 Jan 11;57(1):1-5.

Culpepper WJ, 2nd, Ehrmantraut M, Wallin MT, Flannery K, Bradham DD. Veterans Health Administration multiple sclerosis surveillance registry: the problem of case-finding from administrative databases. J Rehabil Res Dev. 2006 Jan-Feb;43(1):17-24.

Chakravarty EF, Murray ER, Kelman A, Farmer P. Pregnancy outcomes after maternal exposure to rituximab. Blood. 2011 Feb 03;117(5):1499-506.

Charlton BM, Molgaard-Nielsen D, Svanstrom H, Wohlfahrt J, Pasternak B, Melbye M. Maternal use of oral contraceptives and risk of birth defects in Denmark: prospective, nationwide cohort study. BMJ. 2016;352:h6712.

Chastek BJ, Oleen-Burkey M, Lopez-Bresnahan MV. Medical chart validation of an algorithm for identifying multiple sclerosis relapse in healthcare claims. J Med Econ. 2010;13(4):618-25.

Dahl J, Myhr KM, Daltveit AK, Gilhus NE. Pregnancy, delivery and birth outcome in different stages of maternal multiple sclerosis. J Neurol. 2008 May;255(5):623-7.

Dahl J, Myhr KM, Daltveit AK, Hoff JM, Gilhus NE. Pregnancy, delivery, and birth outcome in women with multiple sclerosis. Neurology. 2005 Dec 27;65(12):1961-3.

Danish Data Protection Agency. Introduction to the Danish Data Protection Agency. 2011. Available at: http://www.datatilsynet.dk/english/the-danish-data-protection-agency/introduction-to-the-danish-data-protection-agency/. Accessed July 13, 2017.

Danish Health Authority. Website. 2016. Available at: http://sundhedsstyrelsen.dk/en. Accessed July 13, 2017.

Dolk H, Loane M, Garne E. The prevalence of congenital anomalies in Europe. Adv Exp Med Biol. 2010;686:349-64.

Eltonsy S, Martin B, Ferreira E, Blais L. Systematic procedure for the classification of proven and potential teratogens for use in research. Birth Defects Res A Clin Mol Teratol. 2016 Apr;106(4):285-97.

EMA. Guideline on good pharmacovigilance practices (GVP) Module VI – Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2). European Medicines Agency; 28 July 2017a. Available at: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-good-pharmacovigilance-practices-gvp-module-vi-collection-management-submission-reports_en.pdf. Accessed November 12, 2019.

EMA. Guideline on good pharmacovigilance practices (GVP) Module VIII – Post-authorisation safety studies (Rev 3). European Medicines Agency; 9 October 2017b. Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-module-viii-post-authorisation-safety-studies-rev-3_en.pdf. Accessed November 12, 2019.

ENCePP. ENCePP Checklist for Study Protocols (Revision 4). 2018a. Available at: http://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml. Accessed November 12, 2019.

ENCePP. The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology (Revision 7). EMA/95098/2010. European Medicines Agency; July 2018b. Available at: http://www.encepp.eu/standards_and_guidances/documents/ENCePPGuideonMethStandardsinPE_Rev7.pdf. Accessed November 12, 2019.

ENCePP. The European Union electronic register of post-authorisation studies (EU PAS Register). London: European Network of Centres for Pharmacoepidemiology and Pharmacovigilance; 15 Jul 2016. Available at: http://www.encepp.eu/encepp_studies/indexRegister.shtml. Accessed July 13, 2017.

EUROCAT. Minor anomalies for exclusion. October 14, 2014. Available at: http://www.eurocat-network.eu/content/Section%203.2-%2027_Oct2016.pdf. Accessed September 28, 2017.

EUROCAT. Section 3.3, EUROCAT subgroups of congenital anomalies. In: EUROCAT guide 1.4. 2012. Available at: http://www.eurocat-network.eu/aboutus/datacollection/guidelinesforregistration/malformationcodingguides. Accessed July 13, 2017.

European Commission. Commission implementing regulation (EU) No 520/2012 of 19 June 2012 on the performance of pharmacovigilance activities provided for in Regulation (EC) No 726/2004 of the European Parliament and of the Council and Directive 2001/83/EC of the European Parliament and of the Council. 20 June 2012. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:159:0005:0025:EN:PDF. Accessed July 13, 2017.

Eworuke E, Hampp C, Saidi A, Winterstein AG. An algorithm to identify preterm infants in administrative claims data. Pharmacoepidemiol Drug Saf. 2012 Jun;21(6):640-50.

Flachenecker P, Buckow K, Pugliatti M, Kes VB, Battaglia MA, Boyko A, et al. Multiple sclerosis registries in Europe - results of a systematic survey. Mult Scler. 2014 Oct;20(11):1523-32.

Frischer JM, Bramow S, Dal-Bianco A, Lucchinetti CF, Rauschka H, Schmidbauer M, et al. The relation between inflammation and neurodegeneration in multiple sclerosis brains. Brain. 2009 May;132(Pt 5):1175-89.

Frischer JM, Weigand SD, Guo Y, Kale N, Parisi JE, Pirko I, et al. Clinical and pathological insights into the dynamic nature of the white matter multiple sclerosis plaque. Ann Neurol. 2015 Nov;78(5):710-21.

Gelperin K, Hammad H, Leishear K, Bird ST, Taylor L, Hampp C, et al. A systematic review of pregnancy exposure registries: examination of protocol-specified pregnancy outcomes, target sample size, and comparator selection. Pharmacoepidemiol Drug Saf. 2017 Feb;26(2):208-14.

Genentech Inc. Ocrevus prescribing information. March 2017. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/761053lbl.pdf. Accessed July 13, 2017.

Hauser SL, Bar-Or A, Comi G, Giovannoni G, Hartung HP, Hemmer B, et al. Ocrelizumab versus interferon beta-1a in relapsing multiple sclerosis. N Engl J Med. 2017 Jan 19;376(3):221-34.

HealthCore. About HIRD: HealthCore Integrated Research Database. 2017. Available at: https://www.healthcore.com/database/. Accessed July 13, 2017.

Higuera L, Carlin CS, Anderson S. Adherence to disease-modifying therapies for multiple sclerosis. J Manag Care Spec Pharm. 2016 Dec;22(12):1394-401.

Houtchens MK. Pregnancy and multiple sclerosis. Semin Neurol. 2007 Nov;27(5):434-41.

Hviid A, Melbye M, Pasternak B. Use of selective serotonin reuptake inhibitors during pregnancy and risk of autism. N Engl J Med. 2013 Dec 19;369(25):2406-15.

IBM WH. IBM MarketScan Research Databases for Health Services Researchers. White Paper. IBM Watson Health; April 2019. Available at: https://www.ibm.com/downloads/cas/6KNYVVQ2. Accessed October 31, 2019.

ICMJE. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. International Committee of Medical Journal Editors; December 2018. Available at: http://www.icmje.org/recommendations/. Accessed November 12, 2019.

ICH. Pharmacovigilance planning. E2E. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; 2004. Available at: http://www.ich.org/products/guidelines/efficacy/efficacy-single/article/pharmacovigilance-planning.html. Accessed July 13, 2017.

ISPE. Guidelines for good pharmacoepidemiology practices (GPP). Revision 3. International Society for Pharmacoepidemiology; June 2015. Available at: http://www.pharmacoepi.org/resources/guidelines_08027.cfm. Accessed July 13, 2017.

Johannesdottir SA, Horvath-Puho E, Ehrenstein V, Schmidt M, Pedersen L, Sorensen HT. Existing data sources for clinical epidemiology: the Danish National Database of Reimbursed Prescriptions. Clin Epidemiol. 2012;4:303-13.

Johansson K, Cnattingius S, Naslund I, Roos N, Trolle Lagerros Y, Granath F, et al. Outcomes of pregnancy after bariatric surgery. N Engl J Med. 2015 Feb 26;372(9):814-24.

Kelly VM, Nelson LM, Chakravarty EF. Obstetric outcomes in women with multiple sclerosis and epilepsy. Neurology. 2009 Dec 01;73(22):1831-6.

Kildemoes HW, Sorensen HT, Hallas J. The Danish National Prescription Registry. Scand J Public Health. 2011 Jul;39(7 Suppl):38-41.

Klink DT, van Elburg RM, Schreurs MW, van Well GT. Rituximab administration in third trimester of pregnancy suppresses neonatal B-cell development. Clin Dev Immunol. 2008;2008:271363.

Knudsen LB, Olsen J. The Danish Medical Birth Registry. Dan Med Bull. 1998.

Koch-Henriksen N, Magyari M, Laursen B. Registers of multiple sclerosis in Denmark. Acta Neurol Scand. 2015;132(199):4-10.

Koch-Henriksen N, Stenager E, Laursen B. The use of epidemiological multiple sclerosis registers in research: the Danish MS Registry. Acta Neurol Scand Suppl. 2012(195):7-12.

Li DK, Raebel MA, Cheetham TC, Hansen C, Avalos L, Chen H, et al. Genital herpes and its treatment in relation to preterm delivery. Am J Epidemiol. 2014 Dec 1;180(11):1109-17.

Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sorensen PS, Thompson AJ, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. Neurology. 2014 Jul 15;83(3):278-86.

Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. Scand J Public Health. 2011 Jul;39(7 Suppl):30-3.

MacDonald SC, Cohen JM, Panchaud A, McElrath TF, Huybrechts KF, Hernandez-Diaz S. Identifying pregnancies in insurance claims data: Methods and application to retinoid teratogenic surveillance. Pharmacoepidemiol Drug Saf. 2019 Sep;28(9):1211-21.

MacDorman MF, Gregory EC. Fetal and perinatal mortality: United States, 2013. Natl Vital Stat Rep. 2015 Jul 23;64(8):1-24.

Magyari M, Koch-Henriksen N, Sorensen PS. The Danish Multiple Sclerosis Treatment Register. Clin Epidemiol. 2016;8:549-52.

Margulis AV, Palmsten K, Andrade SE, Charlton RA, Hardy JR, Cooper WO, et al. Beginning and duration of pregnancy in automated health care databases: review of estimation methods and validation results. Pharmacoepidemiol Drug Saf. 2015 Apr;24(4):335-42.

Marrie RA, Yu N, Blanchard J, Leung S, Elliott L. The rising prevalence and changing age distribution of multiple sclerosis in Manitoba. Neurology. 2010 Feb 09;74(6):465-71.

Matthews TJ, MacDorman MF, Thoma ME. Infant mortality statistics from the 2013 period linked birth/infant death data set. Natl Vital Stat Rep. 2015 Aug 06;64(9):1-30.

Milo R, Kahana E. Multiple sclerosis: geoepidemiology, genetics and the environment. Autoimmun Rev. 2010 Mar;9(5):A387-94.

Mines D, Tennis P, Curkendall SM, Li DK, Peterson C, Andrews EB, et al. Topiramate use in pregnancy and the birth prevalence of oral clefts. Pharmacoepidemiol Drug Saf. 2014 Oct;23(10):1017-25.

Molgaard-Nielsen D, Svanstrom H, Melbye M, Hviid A, Pasternak B. Association between use of oral fluconazole during pregnancy and risk of spontaneous abortion and stillbirth. JAMA. 2016 Jan 05;315(1):58-67.

Montalban X, Hauser SL, Kappos L, Arnold DL, Bar-Or A, Comi G, et al. Ocrelizumab versus placebo in primary progressive multiple sclerosis. N Engl J Med. 2017 Jan 19;376(3):209-20.

MSIF. Atlas of MS 2013. Mapping multiple sclerosis around the world. Multiple Sclerosis International Federation; 2013. Available at: https://www.msif.org/wp-content/uploads/2014/09/Atlas-of-MS.pdf. Accessed July 13, 2017.

Mueller BA, Zhang J, Critchlow CW. Birth outcomes and need for hospitalization after delivery among women with multiple sclerosis. Am J Obstet Gynecol. 2002 Mar;186(3):446-52.

NMSS. Medications. National Multiple Sclerosis Society; 2017a. Available at: http://www.nationalmssociety.org/Treating-MS/Medications. Accessed July 13, 2017.

NMSS. Primary progressive MS (PPMS). National Multiple Sclerosis Society; 2017b. Available at: http://www.nationalmssociety.org/What-is-MS/Types-of-MS/Primary-progressive-MS. Accessed July 13, 2017.

NMSS. Types of MS. National Multiple Sclerosis Society; 2017c. Available at: http://www.nationalmssociety.org/What-is-MS/Types-of-MS. Accessed July 13, 2017.

NMSS. Who gets MS? (epidemiology). National Multiple Sclerosis Society; 2017d. Available at: http://www.nationalmssociety.org/What-is-MS/Who-Gets-MS. Accessed July 13, 2017.

Oken E, Kleinman KP, Rich-Edwards J, Gillman MW. A nearly continuous measure of birth weight for gestational age using a United States national reference. BMC Pediatr. 2003 Jul 08;3:6.

Orton SM, Herrera BM, Yee IM, Valdar W, Ramagopalan SV, Sadovnick AD, et al. Sex ratio of multiple sclerosis in Canada: a longitudinal study. Lancet Neurol. 2006 Nov;5(11):932-6.

Palmeira P, Quinello C, Silveira-Lessa AL, Zago CA, Carneiro-Sampaio M. IgG placental transfer in healthy and pathological pregnancies. Clin Dev Immunol. 2012;2012:985646.

Pedersen LH, Petersen OB, Ekelund CK, Pedersen L, Tabor A, Sorensen HT. Linkage between the Danish National Health Service Prescription Database, the Danish Fetal Medicine Database, and other Danish registries as a tool for the study of drug safety in pregnancy. Clinical Epidemiology. 2016;8:91-5.

Rassen JA, Solomon DH, Curtis JR, Herrinton L, Schneeweiss S. Privacy-maintaining propensity score-based pooling of multiple databases applied to a study of biologics. Med Care. 2010 Jun;48(6 Suppl):S83-9.

Rothman KJ. Episheet—spreadsheets for the analysis of epidemiologic data. 29 October 2015. Available at: http://www.krothman.org/episheet.xls. Accessed July 13, 2017.

Sahin L. Pregnancy exposure registries and other post-approval studies: current status and observations. Data on file. Center for Drug Evaluation and Research, Office of New Drugs, FDA; 2014. Available at: http://www.fda.gov/downloads/Drugs/NewsEvents/UCM399660.pdf. Accessed January 26, 2016.

Salzer J, Svenningsson R, Alping P, Novakova L, Bjorck A, Fink K, et al. Rituximab in multiple sclerosis: a retrospective observational study on safety and efficacy. Neurology. 2016 Nov 15;87(20):2074-81.

Schmidt M, Pedersen L, Sorensen HT. The Danish Civil Registration System as a tool in epidemiology. Eur J Epidemiol. 2014 Aug;29(8):541-9.

Schmidt M, Schmidt SA, Sandegaard JL, Ehrenstein V, Pedersen L, Sorensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. Clin Epidemiol. 2015;7:449-90.

Simister NE. Placental transport of immunoglobulin G. Vaccine. 2003 Jul 28;21(24):3365-9.

Stegmann BJ, Carey JC. TORCH Infections. Toxoplasmosis, Other (syphilis, varicella-zoster, parvovirus B19), Rubella, Cytomegalovirus (CMV), and Herpes infections. Curr Womens Health Rep. 2002 Aug;2(4):253-8.

STROBE statement. STROBE checklist. Strengthening the Reporting of Observational Studies in Epidemiology; 2007. Available at: http://www.strobe-statement.org/index.php?id=available-checklists. Accessed July 13, 2017.

Toh S, Reichman ME, Houstoun M, Ding X, Fireman BH, Gravel E, et al. Multivariable confounding adjustment in distributed data networks without sharing of patient-level data. Pharmacoepidemiol Drug Saf. 2013 Nov;22(11):1171-7.

Trojano M, Lucchese G, Graziano G, Taylor BV, Simpson S, Jr., Lepore V, et al. Geographical variations in sex ratio trends over time in multiple sclerosis. PLoS One. 2012;7(10):e48078.

Ventura SJ, Curtin SC, Abma JC, Henshaw SK. Estimated pregnancy rates and rates of pregnancy outcomes for the United States, 1990-2008. Natl Vital Stat Rep. 2012 Jun 20;60(7):1-21.

Vukusic S, Kappos L, Wray S, Bader-Weder S, Buffels R, Masterman D, et al. An update on pregnancy outcomes following ocrelizumab treatment in patients with multiple sclerosis and other autoimmune diseases. Poster P710. Presented at the 7th Joint European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS) and Americas Committee for Treatment and Research in Multiple Sclerosis (ACTRIMS) Meeting; 25-28 October 2017. Paris, France.

Wang FT, Liu N, Ng E, Muller A, D'Ambrosio D, Lindenstrøm E, et al. Challenges of identifying relapses among MS patients in a health plan database. Pharmacoepidemiol Drug Saf. 2015;S1:522-3.

WHO. ATC/DDD index. WHO Collaborating Centre for Drug Statistics Methodology. Norwegian Institute of Public Health; 2017. Available at: https://www.whocc.no/atc_ddd_index/. Accessed July 13, 2017.

Wyszynski DF, Carman WJ, Cantor AB, Graham JM, Jr., Kunz LH, Slavotinek AM, et al. Pregnancy and birth outcomes among women with idiopathic thrombocytopenic purpura. J Pregnancy. 2016;2016:8297407.

Appendix 1
Contact Details for Collaborating Institutions

Collaborating Institutions

[redacted] USA
[redacted]; Senior Scientist [redacted] Principal Investigator [redacted] Project Manager

[redacted] USA:
[redacted] Chief Scientific Officer [redacted] Senior Epidemiologist [redacted] Epidemiologist [redacted] Project Manager

[redacted] Denmark
[redacted] Principal Investigator [redacted] Project Leader [redacted] Senior Epidemiologist [redacted] Senior Biostatistician

Appendix 2
ENCePP Checklist for Study Protocols

ENCePP Checklist for Study Protocols (Revision 4)

Adopted by the ENCePP Steering Group on 15/10/2018

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) welcomes innovative designs and new methods of research. This Checklist has been developed by ENCePP to stimulate consideration of important principles when designing and writing a pharmacoepidemiological or pharmacovigilance study protocol. The Checklist is intended to promote the quality of such studies, not their uniformity. The user is also referred to the ENCePP Guide on Methodological Standards in Pharmacoepidemiology, which reviews and gives direct electronic access to guidance for research in pharmacoepidemiology and pharmacovigilance.

For each question of the Checklist, the investigator should indicate whether or not it has been addressed in the study protocol. If the answer is “Yes”, the section number of the protocol where this issue has been discussed should be specified. It is possible that some questions do not apply to a particular study (for example, in the case of an innovative study design). In this case, the answer ‘N/A’ (Not Applicable) can be checked and the “Comments” field included for each section should be used to explain why. The “Comments” field can also be used to elaborate on a “No” answer.

This Checklist should be included as an Annex by marketing authorisation holders when submitting the protocol of a non-interventional post-authorisation safety study (PASS) to a regulatory authority (see the Guidance on the format and content of the protocol of non-interventional post-authorisation safety studies). The Checklist is a supporting document and does not replace the format of the protocol for PASS presented in the Guidance and Module VIII of the Good pharmacovigilance practices (GVP).

Study title: Multisource Surveillance Study of Pregnancy and Infant Outcomes in Ocrelizumab-Exposed Women With Multiple Sclerosis

EU PAS Register® number:
Study reference number: BA39732 (protocol number)

Appendix 2: ENCePP Checklist for Study Protocols (Cont.)

Section 1: Milestones	Yes	No	N/A	Section Number
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	4, 6
1.1.2 End of data collection[2]	☒	☐	☐	4, 6
1.1.3 Progress report(s)	☒	☐	☐	4, 6, 12
1.1.4 Interim report(s)	☐	☐	☒	
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	6
1.1.6 Final report of study results	☒	☐	☐	6

Comments:

No interim progress reports are planned; ocrelizumab use monitoring reports are planned annually from 2020 through 2028.

Section 2: Research question	Yes	No	N/A	Section Number
2.1 Does the formulation of the research question and objectives clearly explain:				
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)	☒	☐	☐	4, 7
2.1.2 The objective(s) of the study?	☒	☐	☐	4, 8
2.1.3 The target population? (i.e., population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	8
2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☐	☒	
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☒	☐	

Comments:

There are no a priori hypotheses in this study.

Section 3: Study design	Yes	No	N/A	Section Number
3.1 Is the study design described? (e.g., cohort, case-control, cross-sectional, other design)	☒	☐	☐	9.1
3.2 Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	9, 9.1

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

Appendix 2: ENCePP Checklist for Study Protocols (Cont.)

Section 3: Study design		Yes	No	N/A	Section Number
3.3	Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	9.7.2
3.4	Does the protocol specify measure(s) of association? (e.g., relative risk, odds ratio, excess risk, incidence rate ratio, hazard ratio, number needed to harm [NNH])	☒	☐	☐	9.7.4
3.5	Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g., adverse events that will not be collected in case of primary data collection)	☒	☐	☐	11

Comments:

Section 4: Source and study populations		Yes	No	N/A	Section Number
4.1	Is the source population described?	☒	☐	☐	9.2
4.2	Is the planned study population defined in terms of:				
	4.2.1 Study time period	☒	☐	☐	9.2.2
	4.2.2 Age and sex	☒	☐	☐	9.2.1
	4.2.3 Country of origin	☐	☐	☒	
	4.2.4 Disease/indication	☒	☐	☐	9.1, 9.2.1
	4.2.5 Duration of follow-up	☒	☐	☐	9.2.3
4.3	Does the protocol define how the study population will be sampled from the source population? (e.g., event or inclusion/exclusion criteria)	☒	☐	☐	9.2.1

Comments:

Eligibility in the study is independent of country of origin.

Section 5: Exposure definition and measurement		Yes	No	N/A	Section Number
5.1	Does the protocol describe how the study exposure is defined and measured? (e.g., operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	9.3.1
5.2	Does the protocol address the validity of the exposure measurement? (e.g., precision, accuracy, use of validation sub-study)	☒	☐	☐	9.9
5.3	Is exposure categorised according to time windows?	☒	☐	☐	9.3.1
5.4	Is intensity of exposure addressed? (e.g., dose, duration)	☐	☐	☐	

Appendix 2: ENCePP Checklist for Study Protocols (Cont.)

Section 5: Exposure definition and measurement	Yes	No	N/A	Section Number
5.5 Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☒	☐	☐	7, 9.3.1
5.6 Is (are) an appropriate comparator(s) identified?	☒	☐	☐	4, 8, 9

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?	☒	☐	☐	9.3.2
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	9.3.2
6.3 Does the protocol address the validity of outcome measurement? (e.g., precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☒	☐	☐	9.9
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)	☒	☐	☐	9.7.4
7.2 Does the protocol address selection bias? (e.g., healthy user/adherer bias)	☒	☐	☐	9
7.3 Does the protocol address information bias? (e.g., misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	9.9

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, subgroup analyses, anticipated direction of effect)	☐	☒	☐	

Comments:

Effect modification is not considered at this point.

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

Comments:

Section 10: Analysis plan	Yes	No	N/A	Section Number
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	9.7
10.2 Is study size and/or statistical precision estimated?	☒	☐	☐	9.5
10.3 Are descriptive analyses included?	☒	☐	☐	9.7.2
10.4 Are stratified analyses included?	☒	☐	☐	9.7.8
10.5 Does the plan describe methods for analytic control of confounding?	☒	☐	☐	9.7.4

Appendix 2: ENCePP Checklist for Study Protocols (Cont.)

Section 10: Analysis plan	Yes	No	N/A	Section Number
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☒	☐	☐	9.3.2
10.7 Does the plan describe methods for handling missing data?	☒	☐	☐	9.7.5
10.8 Are relevant sensitivity analyses described?	☒	☐	☐	4, 9.7.9

Comments:

Section 11: Data management and quality control	Yes	No	N/A	Section Number
11.1 Does the protocol provide information on data storage? (e.g., software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	9.6
11.2 Are methods of quality assurance described?	☒	☐	☐	9.8
11.3 Is there a system in place for independent review of study results?	☒	☐	☐	9.8

Comments:

Section 12: Limitations	Yes	No	N/A	Section Number
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☒	☐	☐	9
12.1.2 Information bias?	☒	☐	☐	9.9
12.1.3 Residual/unmeasured confounding? (e.g., anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods)	☒	☐	☐	9.9
12.2 Does the protocol discuss study feasibility? (e.g., study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☒	☐	☐	9.5

Comments:

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

Appendix 2: ENCePP Checklist for Study Protocols (Cont.)

Section 13: Ethical issues	**Yes**	**No**	**N/A**	**Section Number**
13.3 Have data protection requirements been described?	☒	☐	☐	9.6, 10

Comments:

There has not been any ethical review procedure yet.

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

Comments:

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

Comments:

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

Date: dd/Month/year

Signature: _____

Appendix 3
Algorithms to Identify Key Study Endpoints

Algorithms to Identify Key Endpoints

Endpoint	ICD-10 Codes	CPT/HCPCS Codes	Window for Outcome Ascertainment	Maternal or Infant Files
Major congenital malformations	Included: Q chapter, D215, D821, D1810, P350, P351, P371. Not included (codes for minor malformations[a]): Q671, Q6740, Q672, Q189, Q670, Q673, Q753, Q674, Q135, Q101, Q102, Q0782, Q752, Q103, Q105, Q1880, Q170, Q173, Q175, Q174, Q171, Q172, Q181, Q179, Q6741, Q3850, Q186, Q382, Q184, Q187, Q185, Q381, Q182, Q180, Q680, Q7400, Q6810, Q845, Q8280, Q653-Q656, Q668, Q669, Q665, Q663, Q666, Q662, Q667, Q664, Q833, Q8252, Q8250, Q825, Q8251, Q7660, Q7662, Q765, Q683, Q684, Q685, Q675, Q7643, Q676, Q6821, Q677, Q678, Q760, Q7671, Q0461, Q270, Q250 if gestational age < 37 weeks, Q2111, Q256 if gestational age < 37 weeks, Q261, Q2541, Q331, Q3310, Q314, Q315, Q320, Q4021, Q4320, Q4381, Q4382, Q401, Q430, Q400, Q633, Q610, Q627, Q5521, Q527, Q525, Q523, Q5520, Q53, Q899, Q950, Q951.	-	Birth through first year of life	Infant
Spontaneous abortion	O02.1 O03	59820 59821	Before gestational week 20	Maternal
Stillbirth	O31.0, P95, Z37.1, Z37.3, Z37.4, Z37.6, Z37.7	-	At or after gestational week 20	Maternal
Small for gestational age	US data sources: 1. O36.51 and O36.59 P05.1 2. Combinations of estimated gestational age at birth (derived from codes Z3A), sex, and codes for birth weight (codes P07) based on US growth standards (Oken et al., 2003) 3. #1 or #2	-	90 days before or after delivery	Either
	Danish Registers: 1. Birth weight < 10th percentile of birth weight by gestational age at birth and sex (primary algorithm) 2. Birth weight < 5th percentile of birth weight by gestational age at	Not applicable	Not applicable	Infant

Endpoint	ICD-10 Codes	CPT/HCPCS Codes	Window for Outcome Ascertainment	Maternal or Infant Files
	birth and sex (secondary algorithm)			
Preterm birth[b]	1. ICD-10 codes mapped from ICD-9 codes retained in the final model of the method developed by Eworuke et al. (2012). Mapping will use the mapping tools developed by the Centers for Medicaid and Medicaid Services.[c] 2. Gestational age at birth < 37 completed weeks using the estimated gestational age at birth (e.g., codes Z3A.x) 3. #1 or #2	Per Table 5 in Eworuke et al.	Birth through 3 months of life	Infant

Note: the occurrence of any of these codes one or more times will be taken to represent the presence of the endpoint.

[a] Some of these codes represent more than one minor malformation.

[b] In Denmark, preterm birth will be identified as births before 37 completed gestational weeks using gestational age at birth as recorded in the Danish Medical Birth Register. In US data, validation efforts will determine which algorithm will be preferred.

[c] For algorithm 1, in addition to ICD-10 and *Current Procedural Terminology* codes, length of stay and death will be used to obtain the prematurity score for each child. Scores > 0.8 will be used to identify preterm infants.

Appendix 4
Characteristics of Study Data Sources

Appendix 4: Characteristics of Study Data Sources (Cont.)

Overview of Characteristics of Study Data Sources

Description	HIRD[a]	DAPI and Optum Research Database[b]	MarketScan Databases[c]	Danish National Registers[d]	Danish MS Register[e]	Danish MS Treatment Register[f]
Year available	2006	1993	1995	1977	1956	1996
Database type	Health care claims	Health care claims	Health care claims	Nationwide health record databases capable of linkage with other databases through a unique personal identification number	Disease-specific register	DMT-treated, disease-specific register
Purpose	Administrative	Administrative	Administrative	Administrative	Research database	Clinical quality monitoring, safety surveillance, and research database
Country	US	US	US	Denmark	Denmark	Denmark
Database size	As of June 2017: 60 million individuals with multiple health plans coverage and over 44 million lives with medical and pharmacy coverage at any point since 2006.	Over 100 million lives since 1993	225 million unique patients since 1995 62.9 million covered lives in the most recent update	5.6 million Nationwide	Nationwide	Nationwide
Representa-tiveness of patients	Commercially insured population	Commercially insured population	Commercially insured population	Total population	All patients with MS	All patients with MS treated with DMTs (mandatory registration)

Appendix 4: Characteristics of Study Data Sources (Cont.)

Description	HIRD[a]	DAPI and Optum Research Database[b]	MarketScan Databases[c]	Danish National Registers[d]	Danish MS Register[e]	Danish MS Treatment Register[f]
Type of health care contact or source of data	Physician, specialist, and emergency room visits, and hospital stays	Inpatient hospital, outpatient hospital, emergency room, physician's office, surgery center, etc.) for virtually all types of services provided, including specialty, preventive and office-based treatments	Inpatient hospital, outpatient hospital, emergency room, physician's office,	Inpatient, outpatient (ambulatory) clinics, or emergency department	1,322 departments of neurology, private neurologists, MS rehabilitation clinics, the Danish MS Treatment Registry, the Danish MS Society and neuropathologists. Two neurologists extract data from medical records Online data collection by the neurologist in the MS clinics and monthly quality-control checks of the data by the registry staff	All Danish departments of neurology through notification at treatment start, and during follow-up at month 3 from start and every 6 months thereafter

Appendix 4: Characteristics of Study Data Sources (Cont.)

Description	HIRD[a]	DAPI and Optum Research Database[b]	MarketScan Databases[c]	Danish National Registers[d]	Danish MS Register[e]	Danish MS Treatment Register[f]
Data on medications	Approved prescription drug and biologic products dispensed in pharmacies, specialty pharmacies, outpatient settings, and hospitals	Pharmacy claims or procedure claims for approved drug and biologic products dispensed in pharmacies or specialty pharmacies or dispensed or administered in outpatient settings and hospitals	Pharmacy claims, claims for mail order prescriptions and specialty pharmacies	Pharmacy-dispensed prescriptions, reimbursed and non-reimbursed	Available as the online data collection platform; serves as data source for the clinical quality database where notification on all patients with MS treated with DMT is mandatory	
Medication information available	Dispensed date, product, strength, days' supply, quantity dispensed, and prescriber specialty	Drug name, drug strength, fill date, quantity dispensed, and days of supply, along with administration date for medications administered by a health care provider	Dispensed date, product, strength, days' supply, quantity dispensed	Formulations strength, DDD (dose unit) and DDDs per package, units per package, date dispensing		
Dose	Based on pharmacy dispensing data	Based on pharmacy dispensing data	Based on pharmacy dispensing data	Dose optional, in free text	Not available for DMTs	
Duration	Based on pharmacy dispensing data	Based on pharmacy dispensing data	Based on pharmacy dispensing data	Based on prescription refills	Not available for DMTs	

Appendix 4: Characteristics of Study Data Sources (Cont.)

Description	HIRD[a]	DAPI and Optum Research Database[b]	MarketScan Databases[c]	Danish National Registers[d]	Danish MS Register[e]	Danish MS Treatment Register[f]
Drug dictionary codes	NDC linked to 14-digit GPI For some drugs and biologic therapy products used in outpatient physician offices, clinics, and inpatient settings: CPT, HCPCS, or, depending on the type of product, ICD-9-CM	NDC	NDC	ATC classification code	ATC classification code	
Diagnosis/ clinical indication	No	No	No	Optional, in free text. Based on proxies	MS disease course, date of onset, date of diagnosis, symptoms at onset, diagnostic accuracy, EDSS	MS disease course, date of onset, date of diagnosis, symptoms at onset, diagnostic accuracy, EDSS, number of relapses in last 12 or 24 months
Outpatient diagnosis	Yes	Yes	Yes	Only outpatient hospital diagnoses	Linkage with Danish National Registries	Linkage with Danish National Registries
Hospital diagnosis	Yes	Yes	Yes	Yes	Linkage with Danish National Registries	Linkage with Danish National Registries
Disease codes	ICD-9-CM, ICD-10-CM	ICD-9-CM, ICD-10-CM	ICD-9-CM, ICD-10-CM	ICD-10-CM	ICD-10	

Appendix 4: Characteristics of Study Data Sources (Cont.)

Description	HIRD[a]	DAPI and Optum Research Database[b]	MarketScan Databases[c]	Danish National Registers[d]	Danish MS Register[e]	Danish MS Treatment Register[f]
Procedure codes	CPT, HCPCS	CPT, HCPCS	CPT, HCPCS	NCSP		
Diagnostic examinations	CPT, HCPCS	CPT, HCPCS	CPT, HCPCS	NCSP (no results)	Evoked potentials, MRI, CSF	Radiological examination (with results)
Laboratory tests	Standard LOINC coding	Procedure and test results available	Procedure		JCV, NAb, Tysabri antibody, CSF analyses	Some tests (with results)
Lifestyle risk factors	No	No	No	Yes	No	No
Source of data for validation	Inpatient and outpatient medical records; patient and provider surveys	Inpatient and outpatient medical records	Not available	Not available	Inpatient and outpatient medical records	Not available
Linkage to other data sources	National mortality, cancer and vaccine registries	National mortality and birth registries, surveys and medical records[g]	EMR, mortality	Danish National Registers	Danish National Registers	Danish National Registers
Approximate time lag (updates per year)	3 months (monthly update)	6 months (semi-annually/quarterly)	6 months (quarterly update)	6 to 8 months (annually)		

ATC = Anatomical Therapeutic Chemical; CPT = Current Procedural Terminology; CSF = cerebrospinal fluid; DAPI = Dynamic Assessment of Pregnancies and Infants; DDD = defined daily dose; DMT = disease-modifying therapy; EDSS = Expanded Disability Status\ Scale; EMR = electronic medical record; GPI = Generic Product Identifier; HCPCS = Healthcare Financing Administration Common Procedure Coding System; ICD-10 = *International Statistical Classification of Diseases and Related Health Problems, 10th Revision*; ICD-10-CM = *International Classification of Diseases, 10th Revision, Clinical Modification*; ICD-9-CM = *International Classification of Diseases, 9th Revision, Clinical Modification*; JCV = JC virus; LOINC = Logical Observation Identifiers Names and Codes; MRI = magnetic resonance imaging; MS = multiple sclerosis; NA = not applicable; NAb = neutralizing antibodies; NCSP = NOMESCO classification of surgical procedures; NDC = National Drug Codes.

Appendix 4: Characteristics of Study Data Sources (Cont.)

[a] Source: HealthCore (2017).

[b] Source: Wyszynski et al. (2016).

[d] Sources: Kildemoes et al. (2011); Lynge et al. (2011); Pedersen et al. (2016); Schmidt et al. (2015).

[e] Sources: Bronnum-Hansen et al. (2011); Flachenecker et al. (2014).

[f] Sources: Flachenecker et al. (2014); Magyari et al. (2016). The Danish MS Treatment Register has been integrated into the Danish MS Register and is no longer available as a stand-alone data source for research.

[g] In limited instances where personally identifiable information is available and appropriate approvals are obtained.