MALBEC
Malignancies in Multiple Sclerosis: Multi-country cohort database studies
French Study

Protocol – SNDS Study

Version V3.0, 24 January 2019

Sponsor: Merck KGaA

Bordeaux PharmacoEpi
Plateforme de recherche en Pharmaco-épidémiologie
CIC Bordeaux CIC1401
INSERM – Université de BORDEAUX – CHU de Bordeaux – Adera
Bâtiment Le Tondu case 41 – 146 rue Léo Saignat – 33076 Bordeaux cedex

GENERAL INFORMATION

Title	Malignancies in Multiple Sclerosis: Multi-country cohort database studies – French Study (MALBEC)
Protocol version identifier	Version 3.0
Date of last version of protocol	24 January 2019
EU PAS register number	EUPAS26535
Active substance	Multiple Sclerosis (MS) modifying drugs including immunomodulatory (IM) medications and Immunosuppressive (IS) medications
Medicinal product	NA
Product reference	NA
Procedure number	NA
Sponsor	Merck, KGaA
Joint PASS	No
Research question and objectives	**Research questions**: - Is the risk of any malignancies in MS patients different to that in the general population? - Does the risk of any malignancies in MS differ overall, by age group and gender as compared to that in the general population? - Is the risk of individual malignancy types in MS patients different overall and by gender to that in the general population? - What is the incidence of any malignancies in patients with MS newly treated with DMD (first ever) according to the DMD received? - Is the exposure to DMDs associated with the risk of any malignancies in MS patients? **Main objective:** to estimate the incidence of any malignancies (including or excluding non-melanoma skin cancer) and by individual malignancy type (if possible), overall and stratified by age, gender and history of malignancy in 1/MS patients in comparison to a sample of non-MS patients from the general population; 2/ untreated patients of the MS cohort and 3/ patients of the MS cohort newly treated with disease modifying drugs (DMD), according to the DMD group. **Secondary objectives:** to characterize the association between DMD treatment exposure and any occurrence of malignancies in

	patients of the MS cohort newly treated with DMDs.
Country of study	France
Author	**Dr Pauline BOSCO-LEVY** Scientific Officer Bordeaux PharmacoEpi (BPE), CIC1401 Université de Bordeaux – CHU de Bordeaux – Adera Bâtiment Le Tondu – case 41 146 rue Léo Saignat – 33076 Bordeaux Cedex, France ☎ +33 (0)5 57 57 92 47 - Fax: +33 (0)5 57 57 47 40 pauline.bosco-levy@u-bordeaux.fr

SPONSOR

Sponsor	**Merck, KGaA** Frankfurter Str. 250, HPC: F135_201, 64293 Darmstadt, Germany www.merckgroup.com

1 TABLE OF CONTENTS

2 LIST OF ABBREVIATIONS

ATC	Anatomical Therapeutic Chemical
CI	Confidence Interval
CNAM-TS	*Caisse Nationale de l'Assurance Maladie des Travailleurs Salariés*
CNIL	*Commission Nationale de l'Informatique et des Libertés* (French data protection commission)
CMU-c	*Couverture Mutuelle Universelle-complémentaire* (100% coverage for socially deprived people)
COPD	Chronic Obstructive Pulmonary Disease
DMD	Disease Modifying Drug
DSR	Disease Risk Score
EDMUS	European Database for MS
EGB	*Echantillon Généraliste de Bénéficiaires*
GA	Glatimater Acetate
GP	General Practitioner
HIV	Human Immunodeficiency Virus
HPV	Human Papillomavirus
HR	Hazard Ratio
IBD	Inflammatory Bowel Disease
ICD-10	International Classification of Disease, 10th revision
INDS	*Institut National des Données de Santé* (National Institute of Health Data)
IFNβ	Interferon beta
IM	Immunomodulators
IR	Incidence Rate
IS	Immunosuppressants
LTD	Long-Term Disease (French list of major chronic diseases with full insurance cover of all claims related to disease)
MS	Multiple Sclerosis
MSA	*Mutualité Sociale Agricole*
NMSC	Non-Melanoma Skin Cancer
OR	Odds Ratio
PMSI	*Programme de Médicalisation des Systèmes d'information*
PY	Person-Year
RR	Relative Risk
RSI	*Régime Social des Indépendants*
SAP	Statistical Analysis Plan
SAR	Statistical Analysis Report
SIR	Standardized Incidence Rate
SNDS	*Système National des Données de Santé* (French National healthcare insurance system database)

3 RESPONSIBLE PARTIES

COORDINATING CENTER

Bordeaux PharmacoEpi (BPE)	**Université de Bordeaux – CHU de Bordeaux – Adera**
INSERM CIC1401	Bâtiment Le Tondu – case 41
	146 rue Léo Saignat – 33076 Bordeaux cedex – France
	☎ +33 (0)5 57 57 46 75 - Fax: +33 (0)5 57 57 47 40
	http://www.pharmacoepi.eu
Pr Nicholas Moore	☎ +33 (0)5 57 57 15 60
President founder	nicholas.moore@u-bordeaux.fr
Dr Patrick Blin	☎ +33 (0)5 57 57 95 63
Chief Scientific Officer	patrick.blin@u-bordeaux.fr
Cécile Droz-Perroteau	☎ +33 (0)5 57 57 47 37
Director	cecile.droz@u-bordeaux.fr
Dr Pauline Bosco-Lévy	☎ +33 (0)5 57 57 92 47
Scientific Officer	pauline.bosco-levy@u-bordeaux.fr
Angela Grelaud-Boussinot	☎ +33 (0)5 57 57 48 52
Project Leader	angela.grelaud@u-bordeaux.fr
Estelle Guiard	☎ +33 (0)5 57 57 47 39
Assistant Project Leader	estelle.guiard@u-bordeaux.fr
Pauline Diez	☎ +33 (0)5 57 57 92 21
Team Leader	pauline.diez@u-bordeaux.fr
Régis Lassalle	☎ +33 (0)5 57 57 47 64
Chief of Biostatistics & Data Management	regis.lassalle@u-bordeaux.fr
Clémentine Lacueille	☎ +33 (0)5 57 57 47 68
Statistician-SAS® Programmer	clementine.lacueille@u-bordeaux.fr

SPONSOR	
Merck, KGaA	Frankfurter Str. 250, HPC: F135_201, 64293 Darmstadt, Germany www.merckgroup.com

4 ABSTRACT

TITLE

Malignancies in Multiple Sclerosis: Multi-country cohort database studies – French Study (MALBEC)

RATIONALE AND BACKGROUND

In the integrated clinical safety database, a numerical imbalance was observed between cladribine and the placebo group. The incidence rate of malignancies in cladribine was 0.37 per 100 person-years while in placebo there were no cases (Leist et al., 2014; Giovannoni et al., 2010).

The pattern of malignancies in the cladribine clinical program in Multiple Sclerosis (MS) did not show an obvious difference compared to the available data on malignancies in the general population, or in MS patients. There is no obvious evidence of an increase of the risk of a particular tumor type in cladribine-treated subjects compared to European reference populations. The effect of age, Immunomodulators (IM) and Immunosuppressants (IS) treatments on the risk of malignancy in patients with MS is currently uncertain.

This study will provide background rates of malignancies in 1) the general population, 2) the overall MS patients' population, 3) and the Disease Modifying Drug (DMD) treated MS patients' population, in order to put post-marketing rates of events of cladribine into perspective.

RESEARCH QUESTION AND OBJECTIVES

The main objective is to estimate the incidence of any malignancies (including or excluding non-melanoma skin cancer) and by individual malignancy type (if possible), overall and stratified by age, gender and history of malignancy in:

- MS patients in comparison to a sample of non-MS patients from the general population;
- untreated patients of the MS cohort;
- patients of the MS cohort newly treated with DMD, according to the DMD group.

The secondary objective is to characterize the association between DMD treatment exposure and any occurrence of malignancies in patients of the MS cohort newly treated with DMDs.

STUDY DESIGN

A cohort study will be first conducted, on a MS cohort including all MS patients identified between 2007 and 2014 from the SNDS database, and with at least 12 months of history prior the inclusion date (pre-inclusion period) and a follow-up until the date of death or the end of the study (31 December 2015), or the first malignancy occurrence, whichever will be the earliest (i.e. 2006-2015 study period). Incidence rates of malignancies estimated in this cohort will be compared to those estimated in non-MS patients of the general population from the EGB database.

Two sub-cohorts will be defined from the initial MS cohort:

- a sub-cohort of untreated MS patients, including all patients without any dispensing of DMD treatment during the study period;
- a sub-cohort of MS patients newly treated with DMDs, including all patients who will received for the first time ever within the study period a DMD treatment without any dispensing of DMD in the 12 months preceding the initiation date. The association between the treatment and the risk of malignancy will be estimated considering no induction period or an induction period varying from the initiation date up to 6 or 12 months during which the occurrence of new

malignancy will be not considered as an event.

In addition, a case-control study nested in the MS newly treated sub-cohort will be conducted to provide estimates on the risk of any malignancies in MS patients according to the cumulative DMD treatment exposure.

POPULATION

MS cohort: All patients extracted from the SNDS database
- aged 18 or older at inclusion date;
- affiliated to the general scheme at inclusion date and during the whole study period;
- with a MS diagnosis, identified between the 01 January 2007 and the 31 December 2014 by using hospitalization for MS, or long term disease status or disability allowance with a diagnosis code of MS, or dispensing of DMD specific to MS;
- with at least 12 months of history prior to the inclusion date.

Untreated MS sub-cohort: patients from the MS cohort, without any dispensing of DMD (specific or non-specific to MS) or any hospitalization for MS treatment administration during the study period.

Newly treated MS sub-cohort: patients from the MS cohort who receive for the first time ever a DMD treatment (specific or non-specific to MS) or a first hospitalization for MS treatment administration during the study period.

Non-MS patients: patients from the general population of the EGB database
- aged 18 or older at 01/01/2007;
- affiliated to the general scheme during the whole study period;
- without any MS diagnosis between the 01 January 2007 and the 31 December 2015.

VARIABLES

The **outcome** of interest will be defined as the first occurrence of malignancy (excluding metastasis) identified during the study period. As defined by the algorithm of the French National Health Insurance, malignancy will be identified by any hospitalization or LTD with ICD-10 codes of cancer or by any specific anti-cancer treatment dispensing, excluding cancer recurrences and secondary tumours.

The **exposure** to DMD:
- immunomodulators (IM): interferon beta-1a, interferon beta-1b, glatiramer acetate, pegylated interferon beta 1-a, or daclizumab;
- immunosuppressant (IS);
  - specific DMD (teriflunomide, fingolimod, dimethyl fumarate, alemtuzumab, or natalizumab);
  - non-specific DMD (methotrexate, cyclophosphamide, mitoxantrone, mycophenolate mofetil, azathioprine, rituximab or tacrolimus).

In the **newly treated MS sub-cohort**, patients will be assigned to either group according to type of the first recorded treatment.

Patients will also be classified according to the number of different IM or IS they used during the follow-up period.

Among **cases and controls of the newly treated MS patients** exposure will be assessed within the exposure window (from the initiation date to the index date) in 2 different ways:
- by the ever use of IM or IS (at least one dispensing within the exposure window);

- by the cumulative duration of use of IM or IS (sum of the theoretical coverage period of all the considered treatment dispensing during the overall exposure window).

The **dates of interest** will be:
- the inclusion date defined as the date of MS diagnosis;
- the initiation date defined as the date of DMD initiation in newly treated MS patients;
- the index date defined as the event date of the cases for the corresponding matched controls of the nested case-control study;
- the date of identification of non-MS patients from the EGB general population: 01/01/2007.

Baseline patients' characteristics (for each cohort):
- demographic variables: gender, age;
- MS variables: time since MS diagnosis (duration of MS).

Medical/Medication history (over the pre-inclusion period including inclusion date for MS patients and untreated MS patients, over the pre-initiation period including initiation date for newly treated MS patients and for the cases and controls of the nested case-control study):
- baseline comorbidities: alcohol use disorders; Chronic Obstructive Pulmonary Disease (COPD); diabetes mellitus; autoimmune disease; infectious disease;
- prior use of female hormones, of corticosteroid treatment: name of drug (brand and class);
- prior DMD treatment for cohort of MS patients only: name of drug (brand and class) and duration of the last one:
- number of visits with any healthcare provider (healthcare use):
- history of relapses.

Characteristics during the follow-up: medical/medications (DMD treatment and corticosteroid treatment); number of visits with any healthcare provider (neurologist, GP, other specialists, nurse and emergency room visits) per year; hospitalizations; occurrence of relapses.

DATA SOURCES

The SNDS database is the nationwide healthcare insurance system database with individual anonymous information on all reimbursed outpatient claims linked to the national hospital-discharge summaries database system and the national death registry, using a unique national pseudonymised identifier. It currently includes 98.8% of the French population, more than 66 million persons from birth (or immigration) to death (or emigration), even if a subject changes occupation or retires.

The EGB is a permanent 1/97th random sample of the SNDS, using a unique national pseudonymised identifier. The EGB is fully representative of the French population in terms of gender, age and mean expenditure reimbursed by individual.

STUDY SIZE

The estimated study size over the study period will be about 100,000 MS subjects for the SNDS database and about 1,000 MS subjects for the EGB database. The estimated study size of non-MS patients from the EGB will be around 380,000 subjects.

DATA ANALYSIS

Population description (after stratification on the history of malignancy)
- a flow chart depicting the number of patients available in the database

satisfying the inclusion criteria;

- description of patients at inclusion date (MS patients and untreated MS patients), at initiation date (newly treated MS patients) or at index date (cases and controls): demographic characteristics and MS variables;
- description of patients during the pre-inclusion period (MS patients and untreated MS patients) or pre-initiation (newly treated MS patients, cases and controls): relapse, medical history, medication and healthcare consumption;
- description of patients during the follow-up period (MS patients, untreated MS patients and newly treated MS patients): relapse, medication and number of visits with any healthcare provider.

Incidence rates

- for MS cohort, untreated MS patients and newly treated MS patients, the estimation of crude annual incidence rates with exact Poisson 95% Confidence Interval (95% CI) and standardized incidence rates (SIR)/100,000 PY, will be calculated and stratified by age, gender and history of malignancy for all malignancies combined excluding Non-Melanoma Skin Cancer (NMSC); all malignancies combined including NMSC; and by individual malignancy type (if enough cases). Among the newly treated MS patients, if enough cases are available, a stratification of results according to the type of IS (specific or not) will be performed. SIRs will be calculated by the application of stratum-specific (sex and age) malignancy incidence rates from the general population (sample of non-MS patients).
- for newly treated MS patients, the analysis will also be performed with an induction period (6 and 12 months).

Association between exposure and outcome

The association between DMD exposure and the risk of malignancy will be assessed in the sub-cohort of newly treated MS patients and among the cases and controls nested in the sub-cohort.

MILESTONES

Protocol SNDS	2018, March
Regulatory aspects and data extraction follow-up with CNAM	2018, March-September
Statistical Analysis Plan	2018, March-September
Data management and statistical analysis	6-8 months after data availability
Final report	8-10 months after data availability

5 AMENDMENTS AND UPDATES

Number	Date	Section of study protocol	Amendment or update	Reason
2.0	04/07/2018	3.1, 3.2.3, 3.4, 3.7.3	Update	The version 2.0 is the result of the update following the favorable opinion with recommendations delivered by the CEREES (*Comité d'Expertise pour les Recherches, les Etudes et les Evaluations dans le domaine de la Santé*) on 22 May 2018.
3.0	12/12/2018		Update	Clarifications according to the feasibility study results performed in the EGB database

6 MILESTONES

Milestones	Planned Date
Protocol SNDS	2018, March
Regulatory aspects and data extraction follow-up with CNAM (SNDS)	2018, March-September
Statistical Analysis Plan	2018, March-September
Data management and statistical analysis	6-8 months after data availability
Final report	8-10 months after data availability

7 RATIONALE AND BACKGROUND

7.1 MALIGNANCIES IN CLADRIBINE CLINICAL TRIAL PROGRAM

In the integrated clinical safety database (*i.e*., the placebo-controlled double-blind and all exposed cohorts), a numerical imbalance was observed between cladribine and the placebo group. The incidence rate of malignancies in cladribine was 0.37 per 100 person-years while in placebo there were no cases (Leist *et al.*, 2014; Giovannoni *et al.*, 2010). The majority of subjects were between 40 and 59 years of age.

Overall there was no obvious pattern or cluster of specific tumor types or locations for either cladribine or placebo (including a total of 37 malignancies, of which 33 occurred in patients treated with cladribine and 4 in patients treated with placebo). The types of malignancies seen in cladribine-treated subjects consisted only of solid tumors (e.g., non-melanoma skin cancer (NMSC), malignant melanoma, colorectal cancer, thyroid cancer, and breast cancer, for the most frequently reported ones). And the pattern of malignancies reported in the program is not what would typically be expected in association with immunosuppression. Notably, there were no cases of leukemia, lymphoma or lymphoproliferative disorders. There was no evidence of a dose effect of cladribine on malignancy occurrence, or any temporal relationship between the start of cladribine therapy and onset of malignancy, or increased in risk over time. Overall there is no conclusive evidence for an increased risk of malignancy with cladribine compared to placebo with a Risk Difference per 100PY of 0.2033 (95% CI: -0.0785, 0.3947) (All Exposed cohort).

The incidence rate for malignancies in the placebo population in the cladribine program is lower than expected when compared to other relapsing MS studies and epidemiological databases. In contrast, the malignancy rate in cladribine was as expected, in line with rates reported from epidemiological observations and not different from rates reported with other disease modifying drugs (DMD). Indeed, an independent meta-analysis of study results from Phase III trials of DMDs used in Relapsing Remitting Multiple Sclerosis (Pakpoor *et al.* 2015) showed that there was no increased risk of malignancy with cladribine compared with other agents. The fact that no malignancies were observed in the placebo group of CLARITY was found to be unique and was significantly different to the malignancy rate of all the other placebo groups from the other trials included in the analysis.

7.2 STUDIES ASSESSING EXPOSURE TO IMMUNOMODULATOR (IM) THERAPY

Several studies have assessed the risk of malignancy in patients with MS treated with DMDs (Kingell *et al.,* 2016; Lebrun *et al.*, 2008; Lebrun *et al.*, 2011; Achiron *et al.*, 2005; Bloomgren *et al.*, 2012; Sandberg-Wollheim *et al.*, 2011). Four of them were prospective cohort studies (Kingwell *et al.*, 2016; Lebrun *et al.*, 2008; Lebrun *et al.*, 2011; Achiron *et al.*, 2005) whereas two were studies based on existing post-marketing surveillance data, clinical trials or claims databases (Bloomgren *et al.*, 2012; Sandberg-Wollheim *et al.*, 2011).

In a recent study, the British Columbia MS database was linked to the provincial Cancer Registry, Vital Statistics death files and Health Registration files (Kingwell *et al.*, 2014). Using a nested case-control design, malignancy cases in MS patients were matched with up to 20 randomly selected MS controls at the date of malignancy diagnosis by sex, age (±5 years) and study entry year using incidence density sampling method. Associations between treatment exposure and overall or specific (breast, colorectal, lung and prostate) malignancies were estimated by conditional logistic regression, adjusted for MS disease duration and age. The study included 5,146 relapsing-onset MS patients who were naive to any MS DMD at the start of follow up and 48,705 person-years of follow up, during which 227 malignancies were diagnosed. Exposure to interferons beta (IFNβ) was not significantly different for cases and controls (OR 1.28; 95% CI 0.87 to 1.88). There was a non-significant trend towards an

increased risk of IFNβ exposure in the breast cancer cases (odds ratio: OR 1.77; 95% CI 0.92 to 3.42), but no evidence of a dose–response effect. Tumour size was similar between IFNβ treated and untreated cases. There was no evidence of an increased malignancy risk with exposure to IFNβ over a 12-year observation period.

The results of the study performed by Kingwell *et al*. are consistent with those from the studies performed in France in which MS patients followed at MS centres were identified through the EDMUS database (Lebrun *et al.*, 2008; Lebrun *et al.*, 2011); In both studies, the findings revealed no increased risk of malignancy with exposure to either the IFNβs or glatimater acetate (GA) compared to untreated patients.

Two other studies also showed consistent results, suggesting that there is no difference in malignancy risk for MS patients compared with the general population (Bloomgren *et al.*, 2012; Sandberg-Wollheim *et al.,* 2011). Bloomgren *et al*. analyzed a large US claims database to determine the risk of malignancy in MS patients treated with intramuscular IFNβ-1a. A total of 95,420 MS cases were identified. Of these, 12,894 were classified as intra-muscular IFNß-1a users. No significant differences in malignancy rates were observed between MS patients who used intramuscular IFNß-1a and non-MS controls, MS patients with no intramuscular IFNß-1a use, or untreated MS patients. The most important limitation of this study is that the mean follow up was approximately 3 years for all groups which is insufficient to detect most of the malignancies. Sandberg-Wollheim *et al*. used pooled safety data from key clinical trials and data from the Merck Serono Global Drug Safety (GDS) database to analyse the risk of malignancy in patients with MS receiving subcutaneous (sc) IFN β-1a (Sandberg-Wollheim *et al.*, 2011). Pooled data from trials showed no statistical difference in the risk of any malignancies between those treated with IFN β-1a and those treated with placebo. In the analysis of the GDS data for confirmed malignancy, the ratio of reported malignancies in patients treated with sc IFN β-1a versus those treated with placebo was 0.09 (95% CI: 0.08-0.10) for any cancer except NMSC suggesting that treatment with sc IFN β-1a does not increase the risk of malignancy in patients with MS.

The findings of the study by Kingwell *et al*. are somewhat in contrast to those from a smaller study from Israel of 1,338 MS patients (15 of whom were cases of malignancy following DMD treatment) which showed borderline associations between non-breast malignancy risk and IFNβ treatment, as well as between breast malignancy risk and GA treatment; however, neither of them reached statistical significance (Achiron *et al.*, 2005). The results of this study have to be considered with caution due to methodological limitations.

7.3 STUDIES ASSESSING EXPOSURE TO IMMUNOSUPRESSANTS (IS) THERAPY

In France, a study was performed using the EDMUS (European Database for MS) data from 9 French hospitals (Lebrun *et al.*, 2008). It included 7,418 patients gathered from January 1995 to 30 June 2006. The adjusted relative risk (RR) of malignancy for patients treated with IS only was 1.96 (95% CI: 0.84-4.96) compared to patients untreated with any DMD while the adjusted RR for those treated with IS+IM compared to untreated patients was 0.54 (95% CI: 0.15-1.91). The risk seemed to be increased in patients treated with immunosuppressive drugs only for their MS although the result was not statistically significant. The risk of malignancy for those receiving only IS was associated with duration of IS exposure: IS/year adjusted RR 1.08 (95% CI: 1.01–1.16) and IS/10 years adjusted RR 1.32 (95% CI: 0.96–1.83). This study contains a considerable number of methodological limitations, thus hampering the robustness of the results.

Lebrun *et al.* (2011) used the EDMUS data from 12 MS centers to update the analysis performed in 2008 (Lebrun *et al.*, 2008). It included 9,269 patients with a history of DMD treatment gathered from January 1995 to 30 June 2009. Among those 18% were treated with IS and 30% were treated with IS+IM. The exposure to new oral drugs (n=151 patients) did not

increase the risk of malignancy (RR=0.5, 95% CI: 0.2-2.23, p=0.15). IS treatment seemed to increase the risk of malignancy, especially skin cancer, as observed in other autoimmune diseases. The relative risk was increased proportionally to the number of courses of IS. In patients that had received 2 different IS (n=990), the risk of malignancy was double although the 95% CI was not significant (RR=2.6, 95% CI: 0.9-6.2). This risk was particularly high if the patient had more than three different IS (RR= 4.8, 95% CI: 1.1-8.7) compared to those untreated. The mean duration of IS treatments was 4.9 ± 4.5 years for patients with malignancy and 3.6 ± 4.5 years for patients with MS and no malignancy. The increase in risk was associated with duration of IS exposure, and was significant for azathioprine (RR=1.9, 95% CI: 1.7–3.4), cyclophosphamide (RR=1.9, 95%CI: 1.3–2.6) and mitoxantrone (RR=1.7, 95% CI: 1.4–3.0).

7.4 STUDIES ASSESSING THE EFFECT OF AGE ON THE RISK OF MALIGNANCY IN MS PATIENTS

Age-specific estimates are rarely provided in studies on the risk of malignancy in MS. Among nine studies that reported the incidence of malignancy in MS, only one population-based study conducted in Denmark reported the relative risk of all malignancies (including NMSC) after first known hospital discharge for MS, by sex and age at discharge (Moller *et al.*, 1991). The incidence of malignancy in patients with MS who were less than 50 years of age at time of entry was significantly higher than in the general population (relative risk: RR=1.7, 95% CI: 1.3-2.2). In addition, the RR of malignancy among MS patients <50 years (in men: RR=1.8, women: RR=1.7) was higher than in those 50 years or older (in men: RR=1.0, women: RR=1.1).

7.5 RISK OF MALIGNANCY IN MULTIPLE SCLEROSIS (MS) PATIENTS

A recent meta-analysis included 38 studies about the epidemiology of malignancy in MS patients (Marrie, 2015). The authors found that in the 38 studies included in the review, estimates for incidence and prevalence varied substantially for most cancers. Among them, 11 studies assessed the incidence of cancer and only 9 of those were population-based. For these 9 studies, the summary estimate for the incidence of any malignancies was 4.39% (2.67–6.1) although it presented a large variability (I^2 statistic=99.8%).

Findings regarding malignancy risk in people with MS as compared to the general population have been inconsistent, and estimates for incidence and prevalence varied substantially for most cancers. The conflicting findings might be due to differences in study design or methods of case ascertainment. The meta-analysis by Marrie *et al*. included 11 comparative studies assessing the incidence of any malignancies in MS patients *vs*. the general population (Marrie *et al.*, 2015). Six of them presented no statistical difference between the 2 groups (Achiron *et al.*, 2005; Midgard *et al.*, 1996, Moller *et al.*, 1991; Nielsen *et al.*, 2006; Sumelahti *et al.*, 2004; Wynn *et al.*, 1990), being one of them was considered of low quality due to study limitations (Christiansen *et al.*, 2010). Amongst the 4 studies that found differences between the 2 groups, only one study in Taiwan reported a higher risk of developing any malignancies amongst MS patients when compared with the control cohort (hazard ratio [HR] 1.85, 95% CI 1.26-2.74). All the other 3 studies reported a decreased risk of developing any malignancies among MS patients when compared with individuals without MS (Bahmanyar *et al.*, 2009; Kingwell *et al.*, 2012; Lebrun *et al.*, 2011).

Regarding the risk of individual malignancy type in MS patients compared to the general population, population-based studies suggested that the risk of meningiomas and urinary tract cancers appeared higher than expected in MS (Marrie *et al.*, 2015). In contrast the risks of pancreatic, ovarian, prostate and testicular cancer were lower than expected (Marrie *et al.*, 2015).

7.6 RATIONALE OF THE STUDY

The pattern of malignancies in the cladribine clinical program in MS (all exposed subjects) did not show an obvious difference compared to the available data on malignancies in the general population, or in MS patients. There is no obvious evidence of an increase of the risk of a particular tumor type in cladribine-treated subjects compared to European reference populations. No dose-relationship could be found.

The effect of age, IM and IS treatments on the risk of malignancy in patients with MS is currently uncertain. MS patients seem to have a similar risk of malignancy than the general population but further studies using external data sources are needed to estimate the risk in MS as compared to general population.

This study will provide background rates of malignancies in 1) the general population, 2) the overall MS patients' population, 3) and the DMDs-treated MS patients' population, in order to put post-marketing rates of events of cladribine into perspective. This study will be conducted in 4 countries (Denmark, the Netherlands, France and United States) according to a common protocol.

This document is the Protocol for the French study and it includes specifically adaptations due to inherent characteristics of French database.

8 RESEARCH QUESTION AND OBJECTIVES

The research questions are listed below:

- Is the risk of any malignancies in MS patients different to that in the general population?
- Does the risk of any malignancies in MS differ overall, by age group and gender as compared to that in the general population?
- Is the risk of individual malignancy types in MS patients different overall and by gender to that in the general population?
- What is the incidence of any malignancies in patients with MS newly treated with DMD (first ever) according to the DMD received?
- Is the exposure to DMDs associated with the risk of any malignancies in MS patients?

The main objective is to estimate the incidence of any malignancies (including or excluding non-melanoma skin cancer) and by individual malignancy type (if possible), overall and stratified by age, gender and history of malignancy in:

- MS patients in comparison to a sample of non-MS patients from the general population;
- untreated patients of the MS cohort;
- patients of the MS cohort newly treated with disease modifying drugs (DMD), according to the DMD group.

The secondary objective is to characterize the association between DMD treatment exposure and any occurrence of malignancies in patients of the MS cohort newly treated with DMDs.

9 RESEARCH METHODS

9.1 STUDY DESIGN

A cohort study will be conducted on a MS cohort, including all patients with a MS diagnosis (as defined in section 9.2.2) identified between 01 January 2007 and 31 December 2014 from the *Système National des Données de Santé* (SNDS) database. The date of the MS diagnosis will be considered as the inclusion date (**Figure 1**). Each patient will have at least 12 months of history prior to the inclusion date and will be followed to the date of death or the end of the study (31 December 2015) or the first malignancy whichever will be the earliest (i.e. 2006-2015 study period). Incidence rates of malignancies estimated in this cohort will be compared to those estimated in non-MS patients of the general population from the EGB (*Echantillon Généraliste de Bénéficiaires)* database.

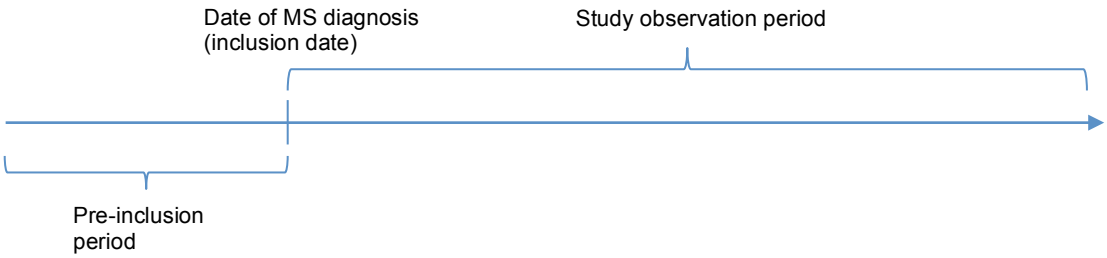

Figure 1. Overall study design in the MS cohort

Two sub-cohorts will be then defined from the initial MS cohort as followed:

- a sub-cohort of untreated MS patients, including all patients without any dispensing of DMD treatment (see Appendix 1) during the study period;
- a sub-cohort of MS patients newly treated with DMDs (first ever), including all patients who will received for the first time ever within the study period (i.e. initiation date) a DMD treatment without any dispensing of DMD in the 12 months preceding the initiation date (i.e. pre-initiation period) (**Figure 2**). The association between the treatment and the risk of malignancy will be estimated considering no induction period or an induction period varying from the initiation date up to 6 or 12 months during which the occurrence of new malignancy will be not considered as an event.

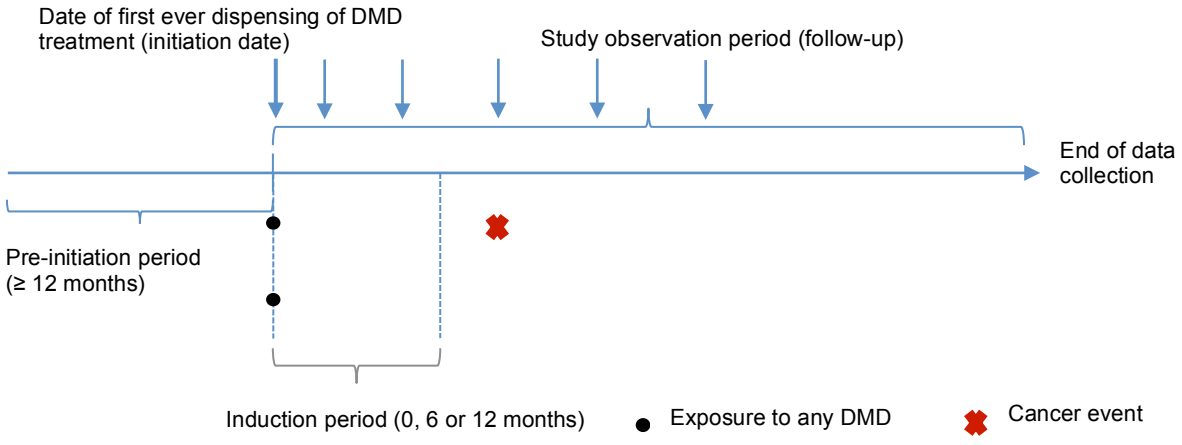

Figure 2. Overall study design in a cohort of MS patients newly treated with DMD

In addition, a case-control study nested in the MS newly treated sub-cohort will be conducted in order to provide estimates on the risk of any malignancies in MS patients considering the exposure as cumulative exposure from initiation date to outcome date. In that **nested case-control study** (Figure 3), cases and controls will be identified in the newly treated MS sub-cohort according to the following definitions:

- **cases** will be all patients with a first occurrence of malignancy during the study period, after an induction period of 6 months following the initiation date (extended to 12 months or reduced to no induction period in sensitivity analysis). The date of the first occurrence of malignancy will correspond to the index date. If more than one cancer of interest is diagnosed in the same patient, only the first diagnosed cancer and its diagnosis date will be considered;
- **controls** will be all patients alive and free of any malignancies over the whole study period. For each control, the index date will be the date of the event (malignancy) of the corresponding matched case.

The exposure window will be defined as the period where the exposure will be measured. It will cover the overall period from the initiation date to the index date.

For each case, up to 6 controls will be matched on age (+/- 5 years), sex, index date, time of the exposure window and a Disease Risk Score (DSR) built using variables collected during the pre-initiation period. Cases will never be considered as potential controls. A same control could be matched to several cases.

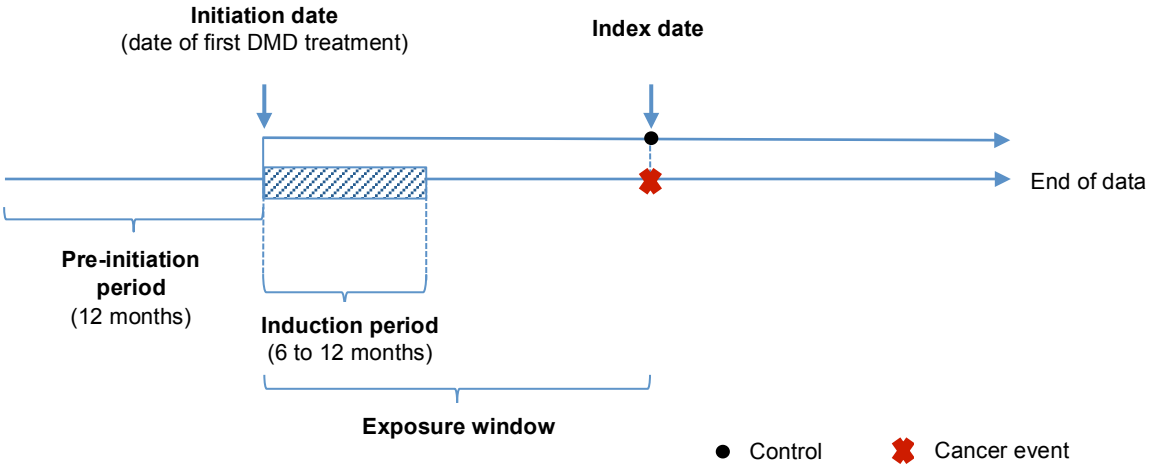

Figure 3. Nested case-control design in the newly treated MS sub-cohort

9.2 SETTING

9.2.1 Treatments of interest

In this study, the treatments of interest are DMDs, including:

- drugs specifically indicated for MS such as immunosuppressants (Aubagio®, Gilenya®, Tecfidera®, Lemtrada® and Tysabri®) or injectable immunomodulators (Avonex®, Betaseron®, Copaxone®, Extavia®, Plegridy®, and Rebif®);
- immunosuppressants non-specifically indicated for MS such as methotrexate, cyclophosphamide, mycophenolate mofetil, aziathropine (Imuran®), mitoxantrone and rituximab.

Each medication will be identified in the French nationwide claims and hospital database (SNDS and EGB) according to the corresponding ATC code (cf. Appendix 1).

9.2.2 Population

9.2.2.1 *MS cohort*

The **MS cohort** will be defined as all patients with a MS diagnosis identified between 01/01/2007 and 31/12/2014 and with at least 12 months of history prior the inclusion date (pre-inclusion period) and no minimum of follow-up. Their data will be extracted from 01/01/2006 to 31/12/2015 (study period). Within the study period, the inclusion date will be the earliest date of MS diagnosis among the following dates:

- date of hospitalization for MS;
- date of hospitalization for a diagnosis specifically related to MS (see Appendix 2);
- starting date of MS LTD or disability;
- date of dispensing for a specific DMD.

For MS patients, the pre-inclusion period will be defined as the 12-month period prior to the inclusion date. Each patient will be followed from the inclusion date until the first date of death, occurrence of malignancy or end of study period, which is set at 31 December 2015.

To be included in **the MS cohort,** patients have to present the following characteristics:

- aged 18 or older at inclusion date;
- affiliated to the general scheme (*Régime Général*, RG) at inclusion date and during the whole study period. Data from the other schemes (*Régime Social des Indépendants* [RSI], *Mutuelle Sociale Agricole* [MSA]), especially those related to the date of death, lack of reliability over the concerned study period;
- with a MS diagnosis, identified between the 01 January 2007 and the 31 December 2014 by one of the following criteria:
  - at least one hospitalization with a diagnosis code of MS (ICD10 code G35) in any position of the claim (main, related or associated diagnosis);
  - a long term disease (LTD) status or a disability allowance with a diagnosis code of MS (ICD10 code G35, see Appendix 2);
  - at least one dispensing of DMD specific to MS: immunosuppressants (Aubagio®, Gilenya®, Tecfidera®, Lemtrada® and Tysabri®), or injectable immunomodulators (Avonex®, Betaseron®, Copaxone®, Extavia®, Plegridy® and Rebif®) (see Appendix 1);
- with at least 12 months of history prior to the inclusion date.

9.2.2.2 *Untreated MS sub-cohort*

The sub-cohort of **untreated MS patients** will be defined as patients from the MS cohort who did not receive any DMD over the study period

In the sub-cohort of untreated MS patients will be included all patients from the MS cohort, without:

- over the study period any dispensing of DMD, whether or not they are specific to MS (see Appendix 1);
- over the study period any hospitalization for MS treatment administration (hospitalization with a treatment administration ICD-10 code Z512 in main diagnosis and an ICD-10 code of MS (G35) or of a diagnosis specifically related to MS (see Appendix 2) in related diagnosis), excluding hospitalizations for MS relapses (as defined in the Appendix 4).

9.2.2.3 *Newly treated MS sub-cohort*

The sub-cohort of **newly treated MS patients** will include a sub-group of patients from the **MS cohort** who will receive a DMD for the first time ever during the study period. The DMD initiation date will be identified at or after the inclusion date. It will correspond to the date of the first DMD dispensing or the first hospitalization for treatment administration whichever comes first. The pre-initiation period will be defined as the 12-month period prior to the initiation date. Each patient will be followed from the initiation date until the first date of death, occurrence of malignancy, or end of study period, which is set at 31 December 2015.

The sub-cohort of newly treated MS patients will include all patients from the MS cohort with:

- a dispensing for a specific or non-specific DMD treatment;
- an hospitalization for a MS treatment administration (hospitalization with a treatment administration ICD-10 code Z512 in main diagnosis and an ICD-10 code of MS (G35) or of a diagnosis specifically related to MS (see Appendix 2) in related diagnosis), excluding hospitalizations for MS relapses (as defined in the Appendix 4).

9.2.2.4 *Non-MS patients*

A selection of **non-MS patients** will be carried out from the general population of the EGB database in order to compare the incidence of malignancy of the MS cohort to that of the general population. Non-MS population will thus include with the following characteristics:

- age ≥ 18 years at 01/01/2007;
- affiliated to the RG during the whole study period;
- without any MS diagnosis between the 01 January 2007 and the 31 December 2014 identified by one of the following criteria:
  - no hospitalization with a diagnosis code of MS (G35 ICD10 code);
  - no long term disease (LTD) status or a disability allowance with a diagnosis code of MS (G35 ICD10 code);
  - no dispensing of DMD specific to MS.

9.2.3 Outcomes

The outcome of interest will be defined as the first occurrence of malignancy (excluding metastasis) identified during the study period. As defined by the algorithm of the French National Health Insurance, malignancy will be identified by any hospitalization or LTD with ICD-10 codes of cancer or by any specific anti-cancer treatment dispensing (see Appendix 3), excluding cancer recurrences and secondary tumours.

Malignancies will be categorized as follows:

- any malignancies, excluding NMSC;
- any malignancies, including NMSC;
- all solid tumours, excluding NMSC;
- all solid tumours, including NMSC;
- all haematological malignancies/lymphoma;
- skin malignancies (basal cell carcinoma, squamous cell carcinoma, NMSC, melanoma of the skin);
- NMSC;
- melanoma of the skin;
- lip, oral cavity cancer;

- nasopharynx cancer;
- other pharynx cancer;
- larynx cancer;
- oesophagus cancer;
- trachea, bronchus and lung cancer;
- brain and nervous system cancer;
- thyroid cancer;
- kidney, renal pelvis and ureter cancer;
- bladder cancer;
- liver and intrahepatic ducts cancer;
- gallbladder and extrahepatic ducts cancer;
- pancreas cancer;
- stomach cancer;
- colorectal cancer, includes anal cancer;
- Hodgkin´s lymphoma;
- non-Hodgkin´s lymphoma;
- multiple myeloma and immunoproliferative disease;
- leukaemia;
- Kaposi sarcoma;
- breast cancer (female only);
- ovarian cancer (female only);
- body of uterus cancer (female only);
- cervix cancer (female only);
- Choriocarcinoma (female only);
- testis cancer (male only);
- prostate cancer (male only);
- eye;
- other female genital organs;
- other male genital organs;
- carcinoma in situ of other sites.

In the unlikely event that too few cases would be identified for some categories of malignancies, events will be grouped into broader categories.

Malignancies will be identified by:

- any hospitalization or LTD with ICD-10 codes of cancer (see Appendix 3);
- any specific anti-cancer treatment dispensing (only for men): L01CD04 (cabazitaxel), L02BX03 (abiratérone acetate), L02BX02 (dégarelix), L02BB01 (flutamide), L02BB02 (nilutamide), L02BB03 (bicalutamide), L02BB04 (enzalutamide), L02AE01 (buséréline), L02AE02 (leuproréline), L02AE03 (goséréline), L02AE04 (triptoréline), G03HA01 (cyprotérone acetate), L01XX11 (estramustine), L02AA01 (diéthylstilbestrol).

Furthermore, if the first malignancy diagnosis is a secondary tumour, the primary malignancy will be identified from all available data during the patient follow-up.

9.3 VARIABLES

9.3.1 Exposure in newly treated MS patients

9.3.1.1 *In the newly treated sub-cohort*

The exposure to DMD will be measured among patients of the **newly treated MS sub-cohort**. A patient will be considered exposed to one of the DMDs of interest if at least one dispensing is recorded in the relevant database. The first DMD treatment dispensed will be considered for selection and the date of the first dispensing or the date of hospitalization for treatment administration after or equals to dispensing date will be defined as the initiation date in this sub-cohort.

Patients will be assigned to the:

- Immunomodulators (IM) treatment group if the first recorded treatment is any IM treatment. A patient can be exposed to one or more IM treatment during follow-up. It includes the following drugs: interferon beta-1a, interferon beta-1b, glatiramer acetate, pegylated interferon beta 1-a, or daclizumab;
- Immunosupressants (IS) treatment group if the first recorded treatment is any IS treatment. A patient can be exposed to one or more IS treatment during follow-up. IS treatment will be also categorized into treatments specific and non specific to MS:
  - specific DMD (teriflunomide, fingolimod, dimethyl fumarate, alemtuzumab, or natalizumab);
  - non-specific DMD (methotrexate, cyclophosphamide, mitoxantrone, mycophenolate mofetil, azathioprine, rituximab or tacrolimus).

The duration exposure of each treatment will be defined as all days exposed to the treatment and calculated as time elapsed between first and last prescription (/dispensing) plus days prescribed and expressed in months.

Treatment discontinuation is considered at the date of the last dispensing plus the days covered by the last dispensing if there is no other dispensing over the remaining follow-up time. The patient will then not initiate a new treatment or switch to another drug in a different class (from IM to IS or from IS to IM).

The period of exposure will then be defined as the duration of treatment from the first dispensing until the first discontinuation or switch (to another treatment group). Patients will be considered as continuously exposed if there is no change in the treatment group over the follow-up, otherwise patients will be flagged as switchers.

Patients will also be classified according to the number of different IM or IS they used during the follow-up period: 1 IM, 2 IM, ≥ 3 IM, 1 IS, 2 IS, ≥ 3 IS, at least one IS and at least one IM.

9.3.1.2 *In cases and controls of the newly treated MS sub-cohort*

The exposure to DMD will also be measured among **cases and controls of the newly treated MS patients** within the exposure window (from the initiation date to the index date) in 2 different ways:

- the primary measure of DMD exposure will be assessed by the ever use of IM or IS defined by the presence of at least one dispensing within the exposure window. Patients will be classified into one of 3 mutually exclusive categories for use of DMDs: exclusive ever use of IS, exclusive ever use of IM and ever use of both IM and IS;
- the second measure of DMD exposure will be assessed by the cumulative duration of use of IM or IS defined as the sum of theoretical coverage period of all the considered treatment dispensing during the overall exposure window.

9.3.2 Other variables

9.3.2.1 *Dates of interest*

The dates of interest will be:

- the inclusion date defined as the date of MS diagnosis;
- the initiation date defined as the date of DMD initiation in newly treated MS patients;
- the index date defined as the event date of the cases for the corresponding matched controls of the nested case-control study;
- the date of identification of non-MS patients from the EGB general population: 01/01/2007.

9.3.2.2 *Demographic variables*

Demographic data will be:

- gender;
- age:
  - at inclusion date for MS patients and untreated MS patients;
  - at initiation date for newly treated MS patients;
  - at index date for the cases and controls of the nested case-control study;
  - at identification date for non-MS patients;

9.3.2.3 *MS variables (at baseline)*

- time since MS diagnosis: duration of MS
  - at inclusion date for MS patients and untreated MS patients;
  - at initiation date for newly treated MS patients and for the cases and controls of the nested case-control study.

9.3.2.4 *Medical/Medication history*

Medical/Medication history will be described over the pre-inclusion period including inclusion date for MS patients and untreated MS patients, over the pre-initiation period including initiation date for newly treated MS patients and for the cases and controls of the nested case-control study. It will concern:

- baseline comorbidities:
  - alcohol use disorders;
  - Chronic Obstructive Pulmonary Disease (COPD);
  - diabetes mellitus;
  - autoimmune disease: thyroid disease, inflammatory bowel disease (IBD), psoriasis, Rheumatoid Arthritis (RA)...;
  - infectious disease: Human Immunodeficiency Virus (HIV);
- prior use of female hormones;
- prior use of corticosteroid treatment: name of drug (brand and class);
- prior DMD treatment for cohort of MS patients only: name of drug (brand and class) and duration of the last one:
  - prior use of IM;
  - prior use of IS;
- number of visits with any healthcare provider (healthcare use):

- neurologist, general practitioner (GP), other specialists and emergency room visits;
- nurse visits;
- new hospitalizations: duration, diagnosis (ICD-10 codes).

Visits with psychologist, psychiatrist, speech therapist, or rehabilitation will be excluded.

- history of relapses (algorithm described in the Appendix 4) (number per patient).

9.3.2.5 *Follow-up*

- medical/medication during the follow-up period for the cohort of MS patients:
  - DMD treatment: name of drug (brand and class), duration at the time of malignancy event;
  - corticosteroid treatment: name of drug (brand and class) and cumulated doses.
- **number of visits** with any healthcare provider (healthcare use) per year:
  - neurologist, GP, other specialists and emergency room visits;
  - nurse visits;

  Visits with psychologist, psychiatrist, speech therapist, or rehabilitation will be excluded;
- **hospitalizations**: duration, diagnoses (ICD-10 codes);
- **occurrence of relapses** (during the follow-up period): number per patient.

9.4 DATA SOURCE

9.4.1 The SNDS

The SNDS database is the nationwide healthcare insurance system database with individual anonymous information on all reimbursed outpatient claims linked to the national hospital-discharge summaries database system (PMSI) and the national death registry, using a unique national pseudonymised identifier. It currently includes 98.8% of the French population nowadays, more than 66 million persons from birth (or immigration) to death (or emigration), even if a subject changes occupation or retires.

Of the 66 million inhabitants in France at the end of 2015, the general scheme covers salaried employees of the private sector and their dependents (*i.e.* about 76% of the population living in France), as well as people covered by *Sections Locales Mutualistes* (SLM), essentially civil servants, employees of territorial collectivities and public hospitals and students, *i.e.* about 11% of the population (Tuppin 2017). The two other main healthcare schemes (*Régime Social des Indépendants* [RSI] - for craftsmen, shopkeepers, liberal professions and their dependents - and *Mutualité Sociale Agricole* [MSA] - for farmers and agricultural workers) cover 11% of the population.

The SNDS contains individual pseudonymised information on (Tuppin 2010, Bezin 2017):

- general characteristics: gender, year of birth, affiliation scheme, area of residence; deprivation status (*Couverture Maladie Universelle complémentaire*, CMU-c);
- date of death for those concerned and very soon cause of death;
- long-term disease (LTD, or ALD in French, and associated ICD-10 codes) with starting and ending date. LTD mainly concerned costly chronic diseases. LTD registration is obtained at the request of a patient's practitioner and validated by the health insurance system physician. Once registered, patients receive full (i.e. 100%) reimbursement for expenditure related to the LTD. The LTD information is specific for the diagnosis (very low risk of false positives), but not sensitive because not all patients with the disease ask to benefit from a LTD;

- outpatient reimbursed healthcare expenditures: visits, medical procedures, nursing acts, physiotherapy, medical imageries, lab tests, drugs, medical devices, transports, sick leaves… with prescriber and professional caregiver information (medical or paramedical specialty, private/public practice), dates (prescription and dispensing), and codes (but not the medical indication nor result);
- hospital-discharge summaries from the PMSI: ICD-10 diagnosis codes (primary, linked and associated diagnosis) for all private and public medical, obstetric and surgery hospitalizations, with the date and duration of hospitalization, medical procedures, and cost coding system, as well as most of very costly drugs. The hospital discharge summary includes the medical unit summaries when the patient is hospitalized successively in several medical units. Primary diagnosis is the health problem that motivated the admission in the hospital. It is determined at hospital discharge. For patients hospitalized successively in several medical units, the primary diagnosis of the hospitalization, as well as all medical unit primary diagnoses, are generally taken into account to define the occurrence of an outcome in a pharmacoepidemiology study. A linked diagnosis can exist only if the primary diagnosis is a care procedure with a code Z of the ICD-10 classification (e.g. chemotherapy session) for a chronic or LTD disease. It indicates the pathology at the origin of the care procedure. As primary diagnosis, is taken into account to define the occurrence of an outcome in a pharmacoepidemiology study. Associated diagnoses are specified if they represent specific healthcare resources. They are mainly underlying chronic diseases. Associated diagnoses can be used to define chronic diseases but are generally not taken into account to define the occurrence of an outcome in a pharmacoepidemiology study (many being false positives for the studied outcome).

Non-hospital data are updated every month with a lag of at least 6 months to have 98% of information uploaded and hospital-discharge summaries yearly at end of Q3 for the previous year. Access to SNDS is regulated and needs approval from the National Institute of Health Data (*Institut National des Données de Santé* - INDS) and the French data protection commission (*Commission Nationale de l'Informatique et des Libertés* - CNIL). The protocol of MALBEC study was approved by the INDS on the 22 May 2018 then authorized by the CNIL on the 10 August 2018.

9.4.2 The EGB

The EGB (*Echantillon Généraliste de Bénéficiaires*) is a permanent 1/97th random sample of the SNDS, using a unique national pseudonymised identifier. Nevertheless, data from psychiatric hospitals and rehabilitation centers are not available in the EGB database. The EGB is fully representative of the French population in terms of gender, age and mean expenditure reimbursed by individual. The Service of Clinical Pharmacology of Bordeaux University has access to the EGB via the INSERM CIC Bordeaux CIC1401 pharmacoepidemiology research unit. As for study based on the SNDS database, an approval from the INDS and the CNIL is needed. The EGB analyses of the MALBEC study protocol were also authorized by the CNIL on the 10 August 2018.

9.5 STUDY SIZE

The study size will not be specified *a priori*, since all patients identified (fulfilling the inclusion criteria) in the selected databases will be analyzed. For the purposes of this analysis, power calculations were not undertaken, as the analysis is not designed under predetermined hypotheses. But *a posteriori* power calculation could be performed for each country.

Nevertheless, related to the results from the clinical trials, the observed cancer incidence occurred in 1.6% (IR = 0.37 cases per 100 person-years [PY]) of the patients. The incidence of a given rare event will follow a Poisson distribution. The unique parameter of the Poisson law is the rate parameter λ, the average number of events observed. If λ_0 is the known background incidence of any event, and δ the additional incidence caused by use of the particular drug under study, then a one-sided test is appropriate. Let's N be the sample size, for a given significance level α and a power 1 – β then the true incidence of the any cases of cancer is given by:

$$N = \frac{[Z_{1-\alpha}\sqrt{\lambda_0} + Z_{1-\beta}\sqrt{(\lambda_0 + \delta)}]^2}{\delta^2}$$

Table 1 gives an indication of the minimum sample size required for δ the additional incidence caused by use of the particular drug under study that may be observed with high probability (> 95%), with a power (1-β) of 80% and for a given significance level α (Machin *et al.*, 2009; Table 1).

Table 1. Minimum sample size that may be observed with high probability (> 95%)

λ	δ	N
0.37	0.01	23,084
0.37	0.02	5,823
0.37	0.03	2,611
0.37	0.04	1,481
0.37	0.05	956

A study size between 1,000 and 23,000 patients per database offers an acceptable level of precision in the different scenarios of available number of patients.

The prevalence of MS in 2012 was estimated to 151 per 100,000 inhabitants using the LTD registration, disability allowance, specific treatments, or hospitalization with MS diagnosis in the SNDS database (Foulon *et al*., 2017). The estimated study size over the study period will be about 100,000 MS subjects for the SNDS database and about 1,000 MS subjects for the EGB database.

The EGB preliminary results were in accordance with this estimation, 1,010 patients without history of cancer were identified in the MS cohort, leading us expecting about 110,000 including 50,000 untreated and 40,000 newly treated (30,000 IM et 10,000 IS) patients in the SNDS database.

The estimated study size of non-MS patients from the EGB will be around 380,000 subjects.

9.6 DATA MANAGEMENT

Database extraction criteria will be described in a Data Extraction Plan (DEP) approved prior to initiating extraction. Extraction of SNDS data will be done by the CNAM.

Data transformation, including decision rules, disease definition, exposure definition, outcomes, risk factors, healthcare resources and calculated variables will be detailed in a statistical analysis plan (SAP).

9.7 DATA ANALYSIS

9.7.1 Generalities

Statistical analysis will be performed using SAS® software (SAS Institute, latest current version, North Carolina, USA). A Statistical Analysis Plan (SAP) will be developed and will be validated before the analysis.

During the regulatory process for SNDS access, a preliminary descriptive analysis will be performed using the EGB database to optimize the SNDS analysis.

The other analyses needed to meet the overall objectives of the study will thus only be performed in the SNDS database.

The following analyses will be performed for:

- the MS cohort (including untreated sub-group);
- the sub-cohort of MS patients newly treated with a DMD treatment;
- the cases and controls of the sub-cohort of MS patients newly treated with a DMD treatment. For each case, up to 6 controls will be matched on age, sex, index date, duration of follow-up and a Disease Risk Score (DRS) built using variables collected during the pre-initiation period. Cases will never be considered as potential controls. A same control could be matched to several cases.

9.7.2 Population description

The selection of study populations for data analysis will be presented in a flow-chart.

The following analyses will be performed for the MS cohort (including untreated sub-group) and newly treated MS patients:

- description of patients at inclusion date (MS patients and untreated MS patients), at initiation date (newly treated MS patients) or at index date (cases and controls): demographic characteristics and MS variables;
- description of patients during the pre-inclusion period (MS patients and untreated MS patients) or pre-initiation (newly treated MS patients, cases and controls): relapse *(not included in the EGB analysis)*, medical history, medication and healthcare consumption;
- description of patients during the follow-up period (MS patients, untreated MS patients and newly treated MS patients): relapse *(not included in the EGB analysis)*, medication and number of visits with any healthcare provider.

9.7.3 Incidence rates

For MS cohort, untreated MS patients and newly treated MS patients, the estimation of crude annual incidence rates with exact Poisson 95% Confidence Interval (95% CI) and standardized incidence rates (SIRs)/100,000 PY, will be calculated and stratified by age, gender and history of malignancy for the following outcomes:

- all malignancies combined excluding NMSC;
- all malignancies combined including NMSC;
- by individual malignancy type (if enough cases).

Among the newly treated MS patients, if enough cases are available, a stratification of results according to the type of IS (specific or not: see in Appendix 1) will be performed.

SIRs will be calculated by the application of stratum-specific (sex and age) malignancy incidence rates from the general population (non-MS patients). The SIR is the ratio of the sum of the observed stratum-specific incidence rates divided by the sum of the expected stratum-specific incidence rates.

For the analysis excluding NMSC (*MS cohort, untreated MS sub-cohort, MS patients newly treated by DMD treatment*): NMSC events will be considered as no event, therefore each patient for whom first event is a NMSC event will have his end of follow-up shifted until the first next non-NMSC event or death or the end of follow-up, whichever came first. For incidence rate estimation, cases with NMSC will be considered in the denominator, and also in the numerator but only if they have a non-NMSC event.

For the analysis with induction period (*MS patients newly treated by DMD treatment*): for each patient, the initiation date will be shifted to the date = [initiation date + induction period + 1 day] *(with induction period = 6 months, 12 months)*. The duration of follow-up and the outcome indicators will be updated according to this shift. Patients with an "end of data collection" date prior to [initiation date + induction period] will be excluded both in the numerator and denominator of the incidence rate estimated in this induction period analysis.

9.7.4 Association between exposure and outcome

The association between DMD exposure and the risk of malignancy will be assessed in the sub-cohort of newly treated MS patients and among the cases and controls nested in the sub-cohort.

9.7.4.1 *Newly treated MS patients*

The association between DMD exposure and the risk of malignancy will be assessed by an intention-to-treat survival analysis using a Cox proportional hazard model. Patients will be assigned in one of each following treatment groups, defined according to the DMDs identified à the initiation date:

- immunomodulators (IM) treatment group if the first recorded treatment is any IM treatment;
- immunosupressants (IS) treatment group if the first recorded treatment is any IS treatment.

Results will be described in term of Hazard Ratio (HR) and its corresponding 95% Confidence Interval (CI), Wald test p-value after adjusting for propensity score. Propensity score is derived from a logistic regression model which estimated for each patient, the probability of being exposed to IM versus IS, conditionally on it observed baseline characteristics. It allows a good control for measured confounders.

9.7.4.2 *Cases and controls nested in newly treated MS patients*

The association between DMD exposure and the risk of malignancy will be assessed using a conditional logistic regression model according to the various definitions of the exposure:

- the ever use of IM or IS, or both;
- the cumulative duration of use of IM or IS.

Results will be described in term of Odds Ratio (OR) and its corresponding 95% Confidence Interval (CI), and Wald test p-value.

9.8 QUALITY CONTROL

The BPE, INSERM CIC1401, has implemented a quality management system for all its activities. CNAM extraction of SNDS data will be validated using the expected population size estimated using the EGB. An independent double programming will be performed for main criteria and analysis, and the results compared for validation. All statistical logs are kept and can be provided. In the case of interim analysis, the database for the interim analysis is locked and kept for ulterior validation if needed. The Statistical Analysis Report (SAR) is included in the final study report.

9.9 LIMITATIONS OF THE RESEARCH METHODS

The first limitation to this study is that such claims database was built for administrative and reimbursement purposes and does not provide clinical data or biological results, including severity or stage of the disease or some risk factors having a confounding effect such as diet, environmental exposures, obesity, alcohol, family history, smoking status, and parity. While this database is fully representative of the French population in terms of gender, age and mean expenditure reimbursed by individual, it appears to not be faithfully representative of cancer cases, especially of those with low prevalence. Moreover, data analysis will be restricted to data from the main healthcare scheme, the only health insurance scheme to have information on the entire study period selected (2006-2015). Patients of other specific healthcare schemes (mainly officials, students, farmers and the self-employed) may differ in terms of their age structure and other socio-demographic characteristics, their exposures and their management and access to healthcare, but they represented only 9% of the overall French population.

Another limitation is that the diagnosis codes recorded of cancer may not be accurate. Certain misclassification bias is so possible, especially in claims databases. To minimize it, an algorithm developed by the CNAM and based on any hospitalization for cancer, any specific anti-cancer treatment dispensing, or LTD registration for cancer will be used to identify cases of malignancy. Cancer recurrence and secondary tumors will be excluded from this definition in order to limit this misclassification bias. Information related to cancer may be missing for some patients, but standardized malignancy rates will be estimated for both MS and non-MS patients using the similar identification algorithm, which will limit this potential bias.

To prevent from wrong or inexact recording of individual factors, either risk factors or the disease being studied, validated code lists of diagnosis mapped on ICD 9 (or ICD-10 for France) international classification will be used.

Another limitation concerns the assessment of drug exposure in the database. Some MS medications (cyclophosphamide and metoxantrone) cannot be identified by their dispensing. Their administration required a one-day hospitalization, which is traceable in the database, but remains difficult to link specifically to one drug. Because of that, some drug exposure may be underestimated and some MS treated patients may have been missed.

The consecutive inclusion of all eligible MS patients into the study will limit possible patient selection bias.

Over a long time horizon, there may be some temporal confounding given significant changes in practice. For instance, better detection techniques can result in higher incidence of malignancy in recent years. In each country, the data quality is considered good throughout the study period and it is not expected different recording of risk factors. Given these reasons, the longest study period in each country that can guarantee good quality data has been selected to allow to observe malignancies that can occur late in time.

MS patients might have an increased frequency of visits with healthcare providers as that of patients in the general population. This might lead to surveillance bias because of an early detection of cancer in MS patients. To address this potential bias, the mean number of visits with any healthcare providers per year will be calculated and compared between MS patients and non-MS patients in the MS cohort, and used as an adjustment factor.

Finally, duration of the MS is difficult to establish in this study. It was approximated by the first dispensing or diagnosis code related to MS within the study period. But patients might have had other dispensing or diagnosis code related to MS before the inclusion date since the study does not focus on new patients. It is thus impossible to determine with certainty when MS was first diagnosed.

9.10 OTHER ASPECTS

None.

10 PROTECTION OF HUMAN SUBJECTS

This project is a database analysis with individual anonymous information for which subject informed consent is not required. Data extraction from the SNDS is regulated and needs approval from National Institute of Health Data (*Institut National des Données de Santé* - INDS) and French data protection commission (*Commission Nationale de l'Informatique et des Libertés* - CNIL).

11 MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

This project is a database analysis using anonymous individual information without any spontaneous reporting. Study outcomes will be reported in aggregate in the final study report, and no individual or expedited reporting is required, according to the EMA Guideline on good pharmacovigilance practices cited above (GVP VI*), as well as the ENCePP Guide on Methodological Standards in Pharmacoepidemiology.

* The latest revision of the Guideline on good pharmacovigilance practices (GVP) Module VI – Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2) from EMA (coming into effect 22 Nov 2017) specifies: *For Non-interventional post-authorisation studies based on secondary use of data (VI.C.1.2.1.2): "The design of such studies is characterised by the secondary use of data previously collected from consumers or healthcare professionals for other purposes. Examples include medical chart reviews (including following-up on data with healthcare professionals), analysis of electronic healthcare records, systematic reviews, meta-analyses. For these studies, the submission of suspected adverse reactions in the form of ICSRs is not required. All adverse events/reactions collected for the study should be recorded and summarised in the interim safety analysis and in the final study report unless the protocol provides for different reporting due to justification*".

12 PLANS FOR DISSEMINATING AND COMMUNATING STUDY RESULTS

This database analysis will be performed by the BPE, INSERM CIC1401, an Academic Research Organization (ARO), for which scientific communication and publication is a major component of its activities. Study methods and results will be submitted to scientific meetings and for publication in international scientific journals.

13 REFERENCES

Achiron A, Barak Y, Gail M, et al. Cancer incidence in multiple sclerosis and effects of immunomodulatory treatments. Breast Cancer Res Treat 2005; 89: 265–270.

Bahmanyar S, Montgomery SM, Hillert J, et al. Cancer risk among patients with multiple sclerosis and their parents. Neurology 2009;72:1170-7

Bezin J, Duong M, Lassalle R, Droz C, Pariente A, Blin P, Moore N. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. Pharmacoepidemiol Drug Saf. 2017 Aug;26(8):954-962

Bloomgren G, Sperling B, Cushing K, et al. Assessment of malignancy risk in patients with multiple sclerosis treated with intramuscular interferon beta-1a: retrospective evaluation using a health insurance claimsdatabase and postmarketing surveillance data. Ther Clin Risk Manage 2012; 8,313–321.

Christiansen CF, Christensen S, Farkas DK, Miret M, Sørensen HT, Pedersen L. Risk of arterial cardiovascular diseases in patients with multiple sclerosis: a population-based cohort study. Neuroepidemiology 2010;35:267-74

ENCePP 2014, European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Code of Conduct, Revision 3, 2014. Available at: http://www.encepp.eu/code_of_conduct/documents/ENCePPCodeofConduct_Rev3.pdf

Foulon S, Maura G, Dalichampt M, Alla F, Debouverie M, Moreau T, Weill A. Prevalence and mortality of patients with multiple sclerosis in France in 2012: a study based on French health insurance data. J Neurol. 2017 Jun;264(6):1185-1192

Giovannoni G, Comi G, Cook S, et al. A placebo-controlled trial of oral cladribine for relapsing multiple sclerosis. New Engl J Med. 2010; 362: 416-26.

Kingwell E, Bajdik C, Phillips N, Zhu F, Oger J, Hashimoto S, Tremlett H. Cancer risk in multiple sclerosis: findings from British Columbia, Canada. Brain. 2012 ;135:2973-9.

Kingwell E, Evans C, Zhu F, Oger J, Hashimoto S, Tremlett H. Assessment of cancer risk with β-interferon treatment for multiple sclerosis. J Neurol Neurosurg Psychiatry. 2014;85:1096-102.

Lebrun C, Debouverie M, Vermersch P, et al. Cancer risk and impact of disease-modifying treatments in patients with multiple sclerosis. Mult Scler 2008;14:399-405.

Lebrun C, Vermersch P, Brassat D, et al. Cancer and multiple sclerosis in the era of disease-modifying treatments. J Neurol. 2011 Jul;258:1304-11.

Leist TP, Comi G, Cree BA, et al. Effect of Oral Cladribine on Time to Conversion to Clinically Definite Multiple Sclerosis in Patients with a First Demyelinating Event (ORACLE MS): A Phase 3 Randomised Trial. Lancet Neurology 2014;13:257-267.

Marrie RA, Reider N, Cohen J, et al. A systematic review of the incidence and prevalence of cancer in multiple sclerosis. Mult Scler 2015;21:294-304

Marrie RA, Cohen J, Stuve O, Trojano M, Sørensen PS, Reingold S, Cutter G, Reider N. A systematic review of the incidence and prevalence of comorbidity in multiple sclerosis: overview. Mult Scler. 2015;21(3):263-81.

Midgard R, Glattre E, Gronning M, et al. Multiple sclerosis and cancer in Norway. A retrospective cohort study. Acta Neurol Scand 1996; 93: 411–415.

Machin D, Michael J. Campbell ML, Beng S, et al. Sample Size Tables for Clinical Studies, Third Edition. Ed. Wiley-Blackwell: London, 2009.

Moller H, Kneller RW, Boice JD Jr, et al. Cancer incidence following hospitalization for multiple sclerosis in Denmark. Acta Neurol Scand 1991; 84: 214–220.Multiple Sclerosis International Federation, 2013. Atlas of MS 2013, London: MSIF.

Nielsen NM, Westergaard T, Frisch M, et al. Type 1 diabetes and multiple sclerosis: A Danish population-based cohort study. Arch Neurol 2006;63:1001-4

Pakpoor J, Disanto G, Altmann DR, Pavitt S, Turner BP, Marta M, Juliusson G, Baker D, Chataway J, Schmierer K. No evidence for higher risk of cancer in patients with multiple sclerosis taking cladribine. Neurol Neuroimmunol Neuroinflamm. 2015;2(6):e158.

Sandberg-Wollheim M, Kornmann G, Bischof D, et al. The risk of malignancy is not increased in patients with multiple sclerosis treated with subcutaneous interferon beta-la: analysis of data from clinical trial and post-marketing surveillance settings. Mult Scler 2011; 17, 431–440.

Sumelahti M-L, Pukkala E, Hakama M. Cancer incidence in multiple sclerosis: a 35-year follow up. Neuroepidemiology 2004; 23: 224–227.

Tuppin P, de Roquefeuil L, Weill A, Ricordeau P, Merlière Y. French nationalhealth insurance information system and the permanent beneficiaries sample. Rev Epidemiol Sante Publique 2010; 58: 286–90.

Tuppin P, Rudant J, Constantinou P, Gastaldi-Ménager C, Rachas A, de Roquefeuil L, et al. Value of a national administrative database to guide public decisions: From the système national d'information interrégimes de l'Assurance Maladie (SNIIRAM) to the système national des données de santé (SNDS) in France. Rev Epidemiol Sante Publique. 2017 Oct;65 Suppl 4:S149–67.

Wynn DR, Rodriguez M, O'Fallon WM, et al. A reappraisal of the epidemiology of multiple sclerosis in Olmsted County, Minnesota. Neurology 1990; 40: 780–786.

14 APPENDICES

Appendix 1. Treatment regimen for DMD

Treatment	*Dose/administration*	*MS specific*	*Availability in EGB/SNDS*	*ATC code*	*Date of 1st Marketing Authorisation in France*
IM treatments (Injectables)					
Avonex® (interferon beta-1a)	*30 mcg intramuscularly (into a large muscle) once weekly*	*Yes*	*Yes*	*L03AB07*	*13/03/1997*
Betaseron® (interferon beta-1b)	*0.25 mg subcutaneously (under the skin) every other day*	*Yes*	*Yes*	*L03AB07*	*30/11/1995*
Copaxone® (glatiramer acetate)	*20 mg subcutaneously every day, or 40 mg subcutaneously 3 times per week*	*Yes*	*Yes*	*L03AX13*	*25/01/2002*
Extavia® (interferon beta-1b)	*0.25 mg subcutaneously every other day*	*Yes*	*Yes*	*L03AB08*	*20/05/2008*
Glatopa™ (glatiramer acetate, generic equivalent of Copaxone 20 mg)	*20 mg subcutaneously every day*	*Yes*	*No*	*A03AX13*	*NC (USA)*
Plegridy® (pegylated interferonbeta-1a)	*125 mcg subcutaneously every 14 days*	*Yes*	*Yes*	*L03AB13*	*18/07/2014*
Rebif® (interferon beta-1a)	*22 mcg or 44 mcg subcutaneously 3 times per week*	*Yes*	*Yes*	*L03AB07*	*04/05/1998*
Zinbryta™ (daclizumab)	150 mg once a month	Yes	No	L04AC01	01/07/2016
IS treatments					
1) Oral					
Aubagio® (teriflunomide)	*7 mg or 14 mg pill once daily*	*Yes*	*Yes*	*L04AA31*	*26/08/2013*
Gilenya® (fingolimod)	*0.5 mg capsule once daily*	*Yes*	*Yes*	*L04AA27*	*17/03/2011*

Tecfidera® (dimethyl fumarate)	*120 mg capsule taken twice daily for one week, followed by 240 mg capsule taken twice daily thereafter*	*Yes*	*Yes*	*N07XX09*	*30/01/2014*
2) Injectables (Intravenous)					
Lemtrada® (alemtuzumab)	*12 mg per day for 5 consecutive days, followed by 12 mg per day on 3 consecutive days one year later*	*Yes*	*No*	*L04AA34*	*12/09/2013*
Mitoxantrone (no brand name available)	*12 mg/m² every 3 months. Lifetime cumulative dose limit of approximately 8–12 doses over 2–3 years (140 mg/m2)*	*No*	*No*	*L01DB07*	*29/10/2003*
Tysabri® (natalizumab)	*300 mg once every 28 days, in an approved infusion facility*	*Yes*	*Yes*	*L04AA23*	*27/06/2006*
3) Other IS used in MS					
Methotrexate	*7.5 mg per week orally*	*No*	*Yes*	*L04AX03*	*NA*
Cyclophosphamide	*Every 4 weeks for one year and every 8 weeks during the second year*	*No*	*Yes*	*L01AA01*	*NA*
Mycophenolate mofetil[α]	*2 grams/day in twice orally daily doses*	*No*	*Yes*	*L04AA06*	*NA*
Imuran® (azathioprine)[β]	*3 mg/kg daily orally*	*No*	*Yes*	*L04AX01*	*NA*
Rituximab	*Intravenous infusion into the bloodstream every two weeks*	*No*	*Yes*	*L01XC02*	*NA*
Tacrolimus[δ]	*0.1 mg/kg/day administered twice daily*	*No*	*Yes*	*L04AD02*	*NA*

α can be used in combination with other DMDs

β can be used in combination with other DMDs, specially interferon beta (1a)

δ can be used in combination with other DMDs, specially with mycophenolate mofetil, cyclophosphamide or Interferon Beta-1b

Appendix 2. – List of ICD-10 codes of interest

Hospitalization	**ICD-10 codes**
Multiple Sclerosis (MS)	G35
Neurological diagnoses specifically related to MS and relapse	
Optic neuritis	H46
Encephalitis, myelitis, encephalomyelitis	G049
Retrobulbar neuritis in diseases classified elsewhere	H481
Other encephalitis, myelitis and encephalomyelitis	G048

Appendix 3. – Algorithm of cancer

The occurrence of the following malignancies will be determined in each database for patients with MS (treated and untreated):

- **Any malignancies, excluding NMSC**: C00-97, + D00-09 but C44, D04, C78, C79 + cancer treatments
- **Any malignancies, including NMSC**: C00-97+ D00-09 but C78, C79 + cancer treatments
- **All solid tumours, excluding NMSC**: C00-80 + C97 + D00-09, but C44, D04, C78, C79 + cancer treatments
- **All solid tumours, including NMSC**: C00-80 + C97 + D00-09 but C78, C79 + cancer treatments
- **All haematological malignancies/lymphoma**: C81-96

ICD-10 codes	Label	Type of cancer
C43	Malignant melanoma of skin	**Skin malignancies (basal cell carcinoma, squamous cell carcinoma, NMSC, melanoma of the skin)**
C44	Other and unspecified malignant neoplasm of skin	
D04	Skin Carcinoma in situ	
D03	Melanoma in situ	
C44	Other and unspecified malignant neoplasm of skin	**Non Melanoma Skin Cancer (NMSC)**
D04	Skin Carcinoma in situ	
C43	Malignant melanoma of skin	**Melanoma of the skin**
D03	Melanoma in situ	
C00	Malignant neoplasm of lip	**Lip, oral cavity cancer**
C01	Malignant neoplasm of base of tongue	
C02	Malignant neoplasm of other and unspecified parts of tongue	
C03	Malignant neoplasm of gum	
C04	Malignant neoplasm of floor of mouth	
C05	Malignant neoplasm of palate	
C06	Malignant neoplasm of other and unspecified parts of mouth	
C07	Malignant neoplasm of parotid gland	
C08	Malignant neoplasm of other and unspecified major salivary glands	
D000	Carcinoma in situ: lip, oral cavity and pharynx	
C11	Malignant neoplasm of nasopharynx	**Nasopharynx cancer**
C09	Malignant neoplasm of tosnil	**Other pharynx cancer**
C10	Malignant neoplasm of oropharynx	
C12	Malignant neoplasm of piriform sinus	
C13	Malignant neoplasm of hypopharynx	
C14	Malignant neoplasm of other and ill-defined sites in the lip, oral cavity and pharynx	
C32	Malignant neoplasm of larynx	**Larynx cancer**
D020	Carcinoma in situ: larynx	
C15	Malignant neoplasm of oesophagus	**Oesophagus cancer**
D001	Carcinoma in situ: Oesophagus	
C33	Malignant neoplasm of trachea	**Trachea, bronchus and lung cancer**
C34	Malignant neoplasm of bronchus and lung	

ICD-10 codes	Label	Type of cancer
D021	Carcinoma in situ: Trachea	
D022	Carcinoma in situ: Bronchus and lung	
D023	Carcinoma in situ: Other parts of respiratory system	
D024	Carcinoma in situ: Respiratory system, unspecified	
C70	Malignant neoplasm of meninges	**Brain and nervous system cancer**
C71	Malignant neoplasm of brain	
C72	Malignant neoplasm of spinal cord, cranial nerves and other parts of central nervous system	
C73	Malignant neoplasm of thyroid gland	**Thyroid cancer**
D093	Carcinoma in situ: Thyroid and other endocrine glands	
D092	Carcinoma in situ: Eye	**Eye**
C64	Malignant neoplasm of kidney, except renal pelvis	**Kidney, renal pelvis and ureter cancer**
C65	Malignant neoplasms of renal pelvis	
C66	Malignant neoplasm of ureter	
C67	Malignant neoplasm of bladder	**Bladder cancer**
D090	Carcinoma in situ: bladder	
D091	Carcinoma in situ: Other and unspecified urinary organs	
C22	Malignant neoplasms of liver and intrahepatic bile ducts	**Liver and intrahepatic ducts cancer**
D015	Carcinoma in situ: Liver, gallbladder and bile ducts	
C23	Malignant neoplasm of gallbladder	**Gallbladder and extrahepatic ducts cancer**
C24	Malignant neoplasm of other and unspecified parts of biliary tract	
C25	Malignant neoplasm of pancreas	**Pancreas cancer**
C16	Malignant neoplasm of stomach	**Stomach cancer**
D002	Carcinoma in situ: Stomach	
C18	Malignant neoplasm of colon	**Colorectal cancer, includes anal cancer**
C19	Malignant neoplasm of rectosigmoid junction	
C20	Malignant neoplasm of rectum	
C21	Malignant neoplasms of anus and anal canal	
D010	Carcinoma in situ: Colon	
D011	Carcinoma in situ: Rectosigmoid junction	
D012	Carcinoma in situ: Rectum	
D013	Carcinoma in situ: Anus and anal canal	
D014	Carcinoma in situ: Other and unspecified parts of intestine	
D017	Carcinoma in situ: Other specified digestive organs	
D019	Carcinoma in situ: Digestive organ, unspecified	
C81	Hodgkin's Disease	**Hodgkin´s lymphoma**
C82	Follicular non-Hodgkin's lymphoma (nodular)	**Non-Hodgkin´s lymphoma**
C83	Diffuse non-Hodgkin's lymphoma	
C84	Peripheral and cutaneous T-cell lymphomas	

ICD-10 codes	Label	Type of cancer
C85	Other and unspecified types of non-Hodgkin's lymphoma	
C96	Other and unspecified malignant neoplasms of lymphoid, haematopoietic and related tissue	
C88	Malignant immunoproliferative diseases	**Multiple myeloma and immunoproliferative disease**
C90	Multiple myeloma and malignant plasma cell neoplasms	
C91	Lymphoid leukemia	**Leukaemia**
C92	Myeloid leukemia	
C93	Monocytic leukemia	
C94	Other leukemias of specified cell type	
C95	Leukemia of unspecified cell type	
C46	Kaposi's Sarcoma	**Kaposi sarcoma**
D097	Carcinoma in situ of other specified sites	**Carcinoma in situ of other sites**
D099	Carcinoma in situ, unspecified	
For female only		
C50	Malignant neoplasm of breast	**Breast cancer**
D05	Carcinoma in situ of breast	
C56	Malignant neoplasm of ovary	**Ovarian cancer**
C54	Malignant neoplasm of corpus uteri	**Body of uterus cancer**
D070	Carcinoma in situ of endometrium	
C53	Malignant neoplasm of cervix uteri	**Cervix cancer**
D06	Carcinoma in situ of cervix uteri	
C58	Malignant neoplasm of placenta	**Choriocarcinoma**
D071	Carcinoma in situ of vulva	**Other female genital organs**
D072	Carcinoma in situ of vagina	
D073	Carcinoma in situ of other and unspecified female genital organs	
For male only		
C62	Malignant neoplasm of testis	**Testis cancer**
C61	Malignant neoplasm of prostate	**Prostate cancer**
D075	Carcinoma in situ of prostate	
D074	Carcinoma in situ of penis	**Other male genital organs**
D076	Carcinoma in situ of Other and unspecified male genital organs	

- **Prostate cancer specific treatment (ATC codes):** L01CD04 (cabazitaxel), L02BX03 (abiratérone acetate), L02BX02 (dégarelix), L02BB01 (flutamide), L02BB02 (nilutamide), L02BB03 (bicalutamide), L02BB04 (enzalutamide), L02AE01 (buséréline), L02AE02 (leuproréline), L02AE03 (goséréline), L02AE04 (triptoréline), G03HA01 (cyprotérone acetate), L01XX11 (estramustine), L02AA01 (diéthylstilbestrol).

Appendix 4. – Definition of relapses

Conditions of relapse	**1**	Hospitalization with at least one of the following diagnoses (ICD-10 codes) G048: Other encephalitis, myelitis et encephalomyelitis, G049: Other encephalitis, myelitis, without specification H46: Optic neuritis, H481: Retrobulbar neuritis
	2	From 3 up to 5 days in hospital for a maximum of 5 days with the 2 following diagnoses (ICD-10 codes) - G35: Multiple sclerosis - Z512: Other forms of chemotherapy
	3	Dispensing of at least 3 grams of corticosteroid (up to 5 grams for the injectable form and up to 6 grams for the oral forms) for a maximum of 5 days
	4	2 days in hospital with a diagnosis code G35 and a simultaneous dispensing of at least 1 gram of corticosteroid (up to 5 grams for the injectable form and up to 6 grams for the oral forms) for a maximum of 5 days
	5	1 day in hospital with a diagnosis code G35 and a simultaneous dispensing of at least 2 grams of corticosteroid (up to 5 grams for the injectable form and up to 6 grams for the oral forms) for a maximum of 5 days

Appendix 5. – ENCePP Check list for the study protocol

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Doc.Ref. EMA/540136/2009

European Network of Centres for Pharmacoepidemiology and Pharmacovigilance

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: Malignancies in Multiple Sclerosis: Multi-country cohort database studies – French Study

EU PAS Register® number: EUPAS26535
Study reference number (if applicable):

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	6 and 9.2.2
1.1.2 End of data collection[2]	☒	☐	☐	6 and 9.2.2
1.1.3 Progress report(s)	☐	☐	☒	
1.1.4 Interim report(s)	☐	☐	☒	
1.1.5 Registration in the EU PAS Register®	☐	☒	☐	

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

30 Churchill Place • Canary Wharf • London E14 5EU • United Kingdom
Telephone +44 (0)20 3660 6000 **Facsimile** +44 (0)20 3660 5555
E-mail: encepp_secretariat@ema.europa.eu **Website** www.encepp.eu

Section 1: Milestones	Yes	No	N/A	Section Number
1.1.6 Final report of study results.	☒	☐	☐	6

Comments:

Section 2: Research question	Yes	No	N/A	Section Number
2.1 Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	8
2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	7.6 and 8
2.1.2 The objective(s) of the study?	☒	☐	☐	8
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	8
2.1.4 Which hypothesis(-es) is (are) to be tested?	☒	☐	☐	8
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

Section 3: Study design	Yes	No	N/A	Section Number
3.1 Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	☒	☐	☐	9.1
3.2 Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	9.1
3.3 Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	9.2.3 and 9.7.3
3.4 Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☒	☐	☐	9.7.3 and 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?	☒	☐	☐	9.2.2 and 9.4
4.2 Is the planned study population defined in terms of:				
4.2.1 Study time period	☒	☐	☐	9.2.2

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.2.2 Age and sex	☒	☐	☐	9.2.2
4.2.3 Country of origin	☒	☐	☐	9.4
4.2.4 Disease/indication	☒	☐	☐	9.2.2
4.2.5 Duration of follow-up	☒	☐	☐	9.2.2
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	9.2.2

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.1 Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	9.2.1 and 9.3.1
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☒	☐	☐	9.7.1
5.3 Is exposure categorised according to time windows?	☒	☐	☐	9.2.2 and 9.3.1
5.4 Is intensity of exposure addressed? (e.g. dose, duration)	☒	☐	☐	9.2.2 and 9.3.1
5.5 Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☐	☒	
5.6 Is (are) (an) appropriate comparator(s) identified?	☒	☐	☐	9.2.2

Comments:

5.5 Not a pharmacokinetic or pharmacodynamic study.

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?	☒	☐	☐	9.2.3
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	9.2.3 and 9.7.3
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)	☒	☐	☐	9.7.3 and 9.7.4
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)	☒	☐	☐	9.3.2

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 and 9.9
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☒	☐	☐	9.9
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	9.7.4 and 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, sub-group analyses, anticipated direction of effect)	☒	☐	☐	9.9

Comments:

8.1 Effect modifiers mentioned in the protocol are not available in the SNDS database

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	9.4
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	9.4
9.1.3 Covariates and other characteristics?	☒	☐	☐	9.4
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	9.4
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)	☒	☐	☐	9.2.1 and 9.4
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	9.2.3 and 9.4

Section 9: Data sources	Yes	No	N/A	Section Number
9.3.3 Covariates and other characteristics?	☒	☐	☐	9.2.2 and 9.4
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☐	☐	☒	

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
10.4 Are stratified analyses included?	☒	☐	☐	9.7
10.5 Does the plan describe methods for analytic control of confounding?	☒	☐	☐	9.7
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☒	☐	☐	9.9
10.7 Does the plan describe methods for handling missing data?	☐	☐	☒	
10.8 Are relevant sensitivity analyses described?	☒	☐	☐	9.7

Comments:

10.5 and 10.7 It will be detailed in the Statistical Analysis Plan developed for the study. 10.6 Validated algorithm used to identify outcomes and validated lists of diagnosis mapped on ICD-10 international classification. 10.7 Study performed from the SNDS database recording all reimbursed claims and hospitalizations without missing values (exhaustive information).

Section 11: Data management and quality control	Yes	No	N/A	Section Number
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	9.8
11.2 Are methods of quality assurance described?	☒	☐	☐	9.8
11.3 Is there a system in place for independent review of study results?	☒	☐	☐	9.8 and 6

Comments:

11.1 This is included specifically in the declaration to the French Data Protection Agency (CNIL). 11.3 Regulatory authorities.

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

Comments:

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	9.4.1 and 10
13.2 Has any outcome of an ethical review procedure been addressed?	☒	☐	☐	9.4.1 and 10
13.3 Have data protection requirements been described?	☒	☐	☐	9.4.1 and 10

Comments:

13. The use of SNDS database is covered by a legal approval by the National Institute of Health Data (INDS) and the French Data Protection Agency (CNIL). No Ethic appoval is needed according to legal status in France.

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: Pauline Bosco-Lévy

Date: 15/01/2019

Signature: [redacted]