



HAL
open science

Calcul du nombre de sujets nécessaires dans les essais cliniques randomisés contrôlés de phase III et IV étudiant les traitements de fond de la sclérose en plaques de forme récurrente-rémittente : une revue systématique

Louis Poncet-Megemont

► To cite this version:

Louis Poncet-Megemont. Calcul du nombre de sujets nécessaires dans les essais cliniques randomisés contrôlés de phase III et IV étudiant les traitements de fond de la sclérose en plaques de forme récurrente-rémittente : une revue systématique. Médecine humaine et pathologie. 2020. dumas-02942764

HAL Id: dumas-02942764

<https://dumas.ccsd.cnrs.fr/dumas-02942764>

Submitted on 18 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ CLERMONT AUVERGNE
UFR DE MÉDECINE ET DES PROFESSIONS PARAMÉDICALES

THÈSE D'EXERCICE
pour le
DIPLÔME D'ÉTAT DE DOCTEUR EN MÉDECINE



PONCET--MEGEMONT, Louis Frédéric

Présentée et soutenue publiquement le 20 Mars 2020



Calcul du nombre de sujets nécessaires dans les essais cliniques randomisés contrôlés de phase III et IV étudiant les traitements de fond de la sclérose en plaques de forme récurrente-rémittente : une revue systématique

Directeur de thèse :

Monsieur MOISSET Xavier, Docteur, CHU de Clermont-Ferrand, Neurologie

Président du jury :

Monsieur CLAVELOU Pierre, Professeur, UFR de Médecine et des Professions paramédicales de Clermont-Ferrand

Membres du jury :

Monsieur GERBAUD Laurent, Professeur, UFR de Médecine et des Professions paramédicales de Clermont-Ferrand

Monsieur RICHARD Ruddy, Professeur, UFR de Médecine et des Professions paramédicales de Clermont-Ferrand

Monsieur PEREIRA Bruno, Biostatisticien, CHU de Clermont-Ferrand, Direction de la Recherche Clinique et de l'Innovation



UNIVERSITE CLERMONT AUVERGNE

PRESIDENTS HONORAIRES
UNIVERSITE D'AUVERGNE

: **JOYON** Louis
: **DOLY** Michel
: **TURPIN** Dominique
: **VEYRE** Annie
: **DULBECCO** Philippe
: **ESCHALIER** Alain

PRESIDENTS HONORAIRES
UNIVERSITE BLAISE PASCAL

: **CABANES** Pierre
: **FONTAINE** Jacques
: **BOUTIN** Christian
: **MONTEIL** Jean-Marc
: **ODOUARD** Albert
: **LAVIGNOTTE** Nadine

PRESIDENT DE L'UNIVERSITE et
PRESIDENT DU CONSEIL ACADEMIQUE PLENIER
PRESIDENT DU CONSEIL ACADEMIQUE RESTREINT
VICE-PRESIDENT DU CONSEIL D'ADMINISTRATION
VICE-PRESIDENT DE LA COMMISSION DE LA RECHERCHE
VICE PRESIDENTE DE LA COMMISSION DE LA
FORMATION ET DE LA VIE UNIVERSITAIRE
DIRECTEUR GENERAL DES SERVICES

: **BERNARD** Mathias
: **DEQUIEDT** Vianney
: **WILLIAMS** Benjamin
: **HENRARD** Pierre

: **PEYRARD** Françoise
: **PAQUIS** François



UFR DE MEDECINE ET DES PROFESSIONS PARAMEDICALES

DOYENS HONORAIRES

: **DETEIX** Patrice
: **CHAZAL** Jean

DOYEN
RESPONSABLE ADMINISTRATIVE

: **CLAVELOU** Pierre
: **ROBERT** Gaëlle

LISTE DU PERSONNEL ENSEIGNANT

PROFESSEURS HONORAIRES :

MM. BACIN Franck - BEGUE René-Jean - BOUCHER Daniel - BOURGES Michel - BUSSIÈRE Jean-Louis - CANO Noël - CASSAGNES Jean - CATILINA Pierre - CHABANNES Jacques - CHAZAL Jean - CHIPPONI Jacques - CHOLLET Philippe - COUDERT Jean - DASTUGUE Bernard - DAUPLAT Jacques - DECHELOTTE Pierre - DEMEOCQ François - DE RIBEROLLES Charles - ESCANDE Georges - Mme FONCK Yvette - MM. GENTOU Claude - GLANDDIER Gérard - Mme GLANDDIER Phyllis - M. JACQUETIN Bernard - Mme LAVARENNE Jeanine - MM. LAVERAN Henri - LESOURD Bruno - LEVAI Jean-Paul - MAGE Gérard - MALPUECH Georges - MARCHEIX Jean-Claude - MICHEL Jean-Luc - MOLINA Claude - MONDIE Jean-Michel - PERI Georges - PETIT Georges - PHILIPPE Pierre - PLAGNE Robert - PLANCHE Roger - PONSONNAILLE Jean - REY Michel - Mme RIGAL Danièle - MM. ROZAN Raymond - SCHOEFFLER Pierre - SIROT Jacques - SOUTEYRAND Pierre - TANGUY Alain - TERVER Sylvain - THIEBLOT Philippe - TOURNILHAC Michel - VANNEUVILLE Guy - VIALLET Jean-François - Mlle VEYRE Annie

PROFESSEURS EMERITES :

MM. - BEYTOUT Jean - BOITEUX Jean-Paul - BOMMELAER Gilles - CHAMOIX Alain - DETEIX Patrice - DUBRAY Claude - ESCHALIER Alain - IRTHUM Bernard - KEMENY Jean-Louis - LABBE André - Mme LAFEUILLE Hélène - MM. LEMERY Didier - LUSSON Jean-René - RIBAL Jean-Pierre

PROFESSEURS DES UNIVERSITES-PRATICIENS HOSPITALIERS

PROFESSEURS DE CLASSE EXCEPTIONNELLE

M. VAGO Philippe	Histologie-Embryologie Cytogénétique
M. AUMAITRE Olivier	Médecine Interne
M. LABBE André	Pédiatrie
M. AVAN Paul	Biophysique et Traitement de l'Image
M. DURIF Franck	Neurologie
M. BOIRE Jean-Yves	Biostatistiques, Informatique Médicale et Technologies de Communication
M. BOYER Louis	Radiologie et Imagerie Médicale option Clinique
M. POULY Jean-Luc	Gynécologie et Obstétrique
M. CANIS Michel	Gynécologie-Obstétrique
Mme PENAULT-LLORCA Frédérique	Anatomie et Cytologie Pathologiques
M. BAZIN Jean-Etienne	Anesthésiologie et Réanimation Chirurgicale
M. BIGNON Yves Jean	Cancérologie option Biologique
M. BOIRIE Yves	Nutrition Humaine
M. CLAVELOU Pierre	Neurologie
M. DUBRAY Claude	Pharmacologie Clinique

M.	GILAIN Laurent	O.R.L.
M.	LEMAIRE Jean-Jacques	Neurochirurgie
M.	CAMILLERI Lionel	Chirurgie Thoracique et Cardio-Vasculaire
M.	DAPOIGNY Michel	Gastro-Entérologie
M.	LLORCA Pierre-Michel	Psychiatrie d'Adultes
M.	PEZET Denis	Chirurgie Digestive
M.	SOUWEINE Bertrand	Réanimation Médicale
M.	BOISGARD Stéphane	Chirurgie Orthopédique et Traumatologie
Mme	DUCLOS Martine	Physiologie
M.	SCHMIDT Jeannot	Thérapeutique
M.	BERGER Marc	Hématologie
M.	GARCIER Jean-Marc	Anatomie-Radiologie et Imagerie Médicale
M.	ROSSET Eugénio	Chirurgie Vasculaire
M.	SOUBRIER Martin	Rhumatologie

**PROFESSEURS DE
1ère CLASSE**

M.	CAILLAUD Denis	Pneumo-physiologie
M.	VERRELLE Pierre	Radiothérapie option Clinique
M.	CITRON Bernard	Cardiologie et Maladies Vasculaires
M.	D'INCAN Michel	Dermatologie -Vénérologie
Mme	JALENQUES Isabelle	Psychiatrie d'Adultes
Mlle	BARTHELEMY Isabelle	Chirurgie Maxillo-Faciale
M.	GERBAUD Laurent	Epidémiologie, Economie de la Santé et Prévention
M.	TAUVERON Igor	Endocrinologie et Maladies Métaboliques
M.	MOM Thierry	Oto-Rhino-Laryngologie
M.	RICHARD Ruddy	Physiologie
M.	RUIVARD Marc	Médecine Interne
M.	SAPIN Vincent	Biochimie et Biologie Moléculaire
M.	BAY Jacques-Olivier	Cancérologie
M.	COUDEYRE Emmanuel	Médecine Physique et de Réadaptation
Mme	GODFRAIND Catherine	Anatomie et Cytologie Pathologiques
M.	ABERGEL Armando	Hépatologie
M.	LAURICHESSE Henri	Maladies Infectieuses et Tropicales
M.	TOURNILHAC Olivier	Hématologie
M.	CHIAMBARETTA Frédéric	Ophthalmologie
M.	FILAIRE Marc	Anatomie – Chirurgie Thoracique et Cardio-Vasculaire
M.	GALLOT Denis	Gynécologie-Obstétrique
M.	GUY Laurent	Urologie
M.	TRAORE Ousmane	Hygiène Hospitalière
M.	ANDRE Marc	Médecine Interne
M.	BONNET Richard	Bactériologie, Virologie
M.	CACHIN Florent	Biophysique et Médecine Nucléaire
M.	COSTES Frédéric	Physiologie
M.	FUTIER Emmanuel	Anesthésiologie-Réanimation
Mme	HENG Anne-Elisabeth	Néphrologie
M.	MOTREFF Pascal	Cardiologie
Mme	PICKERING Gisèle	Pharmacologie Clinique
M.	RABISCHONG Benoît	Gynécologie Obstétrique

**PROFESSEURS DE
2ème CLASSE**

Mme CREVEAUX Isabelle	Biochimie et Biologie Moléculaire
M. FAICT Thierry	Médecine Légale et Droit de la Santé
Mme KANOLD LASTAWIECKA Justyna	Pédiatrie
M. TCHIRKOV Andréï	Cytologie et Histologie
M. CORNELIS François	Génétique
M. DESCAMPS Stéphane	Chirurgie Orthopédique et Traumatologique
M. POMEL Christophe	Cancérologie – Chirurgie Générale
M. CANAVESE Fédérico	Chirurgie Infantile
M. LESENS Olivier	Maladies Infectieuses et Tropicales
M. AUTHIER Nicolas	Pharmacologie Médicale
M. BROUSSE Georges	Psychiatrie Adultes/Addictologie
M. BUC Emmanuel	Chirurgie Digestive
M. CHABROT Pascal	Radiologie et Imagerie Médicale
M. LAUTRETTE Alexandre	Néphrologie Réanimation Médicale
M. AZARNOUSH Kasra	Chirurgie Thoracique et Cardiovasculaire
Mme BRUGNON Florence	Biologie et Médecine du Développement et de la Reproduction
Mme HENQUELL Cécile	Bactériologie Virologie
M. ESCHALIER Romain	Cardiologie
M. MERLIN Etienne	Pédiatrie
Mme TOURNADRE Anne	Rhumatologie
M. DURANDO Xavier	Cancérologie
M. DUTHEIL Frédéric	Médecine et Santé au Travail
Mme FANTINI Maria Livia	Neurologie
M. SAKKA Laurent	Anatomie – Neurochirurgie
M. BOURDEL Nicolas	Gynécologie-Obstétrique
M. GUIEZE Romain	Hématologie
M. POINCLOUX Laurent	Gastroentérologie
M. SOUTEYRAND Géraud	Cardiologie
M. EVRARD Bertrand	Immunologie
M. POIRIER Philippe	Parasitologie et Mycologie

PROFESSEURS DES UNIVERSITES

M. CLEMENT Gilles	Médecine Générale
Mme MALPUECH-BRUGERE Corinne	Nutrition Humaine
M. VORILHON Philippe	Médecine Générale

PROFESSEURS ASSOCIES DES UNIVERSITES

Mme BOTTET-MAULOUBIER Anne	Médecine Générale
M. CAMBON Benoît	Médecine Générale
M. TANGUY Gilles	Médecine Générale

MAITRES DE CONFERENCES DES UNIVERSITES - PRATICIENS HOSPITALIERS

MAITRES DE CONFERENCES HORS CLASSE

Mme CHAMBON Martine	Bactériologie Virologie
Mme BOUTELOUP Corinne	Nutrition

MAITRES DE CONFERENCES DE 1ère CLASSE

M. MORVAN Daniel	Biophysique et Traitement de l'Image
Mlle GOUMY Carole	Cytologie et Histologie, Cytogénétique
Mme FOGLI Anne	Biochimie Biologie Moléculaire
Mlle GOUAS Laetitia	Cytologie et Histologie, Cytogénétique
M. MARCEAU Geoffroy	Biochimie Biologie Moléculaire
Mme MINET-QUINARD Régine	Biochimie Biologie Moléculaire
M. ROBIN Frédéric	Bactériologie
Mlle VERONESE Lauren	Cytologie et Histologie, Cytogénétique
M. DELMAS Julien	Bactériologie
Mlle MIRAND Audrey	Bactériologie Virologie
M. OUCHCHANE Lemlih	Biostatistiques, Informatique Médicale et Technologies de Communication
M. LIBERT Frédéric	Pharmacologie Médicale
Mlle COSTE Karen	Pédiatrie
Mlle AUMERAN Claire	Hygiène Hospitalière
Mme CASSAGNES Lucie	Radiologie et Imagerie Médicale
M. LEBRETON Aurélien	Hématologie
M. BUISSON Anthony	Gastroentérologie

MAITRES DE CONFERENCES DE 2ème CLASSE

Mme PONS Hanaë	Biologie et Médecine du Développement et de la Reproduction
M. JABAUDON-GANDET Matthieu	Anesthésiologie – Réanimation Chirurgicale
M. BOUVIER Damien	Biochimie et Biologie Moléculaire
M. COLL Guillaume	Neurochirurgie
Mme SARRET Catherine	Pédiatrie
M. MAQDASY Salwan	Endocrinologie, Diabète et Maladies Métaboliques
Mme NOURRISSON Céline	Parasitologie - Mycologie

MAITRES DE CONFERENCES DES UNIVERSITES

Mme VAURS-BARRIERE Catherine	Biochimie Biologie Moléculaire
M. BAILLY Jean-Luc	Bactériologie Virologie
Mlle AUBEL Corinne	Oncologie Moléculaire
M. BLANCHON Loïc	Biochimie Biologie Moléculaire
Mlle GUILLET Christelle	Nutrition Humaine
M. BIDET Yannick	Oncogénétique
M. MARCHAND Fabien	Pharmacologie Médicale
M. DALMASSO Guillaume	Bactériologie
M. SOLER Cédric	Biochimie Biologie Moléculaire
M. GIRAUDET Fabrice	Biophysique et Traitement de l'Image
Mme VAILLANT-ROUSSEL Hélène	Médecine Générale
Mme LAPORTE Catherine	Médecine Générale
M. LOLIGNIER Stéphane	Neurosciences – Neuropharmacologie
Mme MARTEIL Gaëlle	Biologie de la Reproduction
M. PINEL Alexandre	Nutrition Humaine
M. PIZON Frank	Santé Publique

MAITRES DE CONFERENCES ASSOCIES DES UNIVERSITES

M. BERNARD Pierre	Médecine Générale
Mme ESCHALIER Bénédicte	Médecine Générale
Mme RICHARD Amélie	Médecine Générale
M. TESSIERES Frédéric	Médecine Générale

« Je ne perds jamais, soit je gagne, soit j'apprends. »

Nelson Mandela

REMERCIEMENTS ET DEDICACES

A mon directeur de thèse, Monsieur le Docteur Xavier MOISSET,

Merci d'avoir accepté de diriger ce travail de thèse. Celui-ci n'aurait jamais pu aboutir sans tes conseils avisés. Merci pour ta disponibilité, ton sérieux et ton humanisme qui sont un exemple à suivre pour tout jeune neurologue.

Au président du jury, Monsieur le Professeur Pierre CLAVELOU,

Merci de m'avoir accordé votre confiance pour effectuer ce travail de thèse. Je vous exprime toute ma gratitude pour votre accompagnement durant ces années d'internat.

Aux membres du jury

Monsieur le Professeur Laurent GERBAUD,

Merci d'avoir accepté de participer à mon jury de thèse. Je suis honoré de pouvoir bénéficier de votre expertise afin de juger ce travail.

Monsieur le Professeur Ruddy RICHARD,

Merci d'avoir accepté d'évaluer mon travail de thèse. Votre expérience de la recherche clinique est un réel atout pour ce jury.

Monsieur Bruno PEREIRA,

Tu as su changer ma vision du monde des biostatistiques. Merci pour tout le travail effectué à mes côtés et pour tes rendez-vous matinaux.

A Amélie, mes parents, mon frère et mes amis,

Merci de m'avoir soutenu pendant toutes ces années, et d'être toujours présents au quotidien.

TABLE DES MATIERES

LISTE DES ABREVIATIONS	11
INTRODUCTION	12
1. Contexte scientifique	13
<i>1.1. Importance de la qualité des publications scientifiques</i>	13
<i>1.2. Calcul du nombre de sujets nécessaires</i>	14
<i>1.3. Généralités sur la Sclérose En Plaques</i>	15
<i>1.4. Calcul du NSN basé sur le taux annualisé de poussées</i>	16
2. Hypothèses de travail et objectifs	18
3. Références bibliographiques	19
ARTICLE	20
CONCLUSION	45
ANNEXES	47
SERMENT D'HIPPOCRATE	52

LISTE DES ABREVIATIONS

CJP : Critère de Jugement Principal

NSN : Nombre de Sujets Nécessaires

PP : Progressive Primaire

RR : Récurrente-Rémittente

SEP : Sclérose En Plaques

SP : Secondairement Progressive

TAP : Taux Annualisé de Poussées

INTRODUCTION

1. Contexte scientifique

1.1. Importance de la qualité des publications scientifiques

La prise de décision médicale, l'établissement de recommandations de bonnes pratiques cliniques et le remboursement de nouvelles thérapies reposent actuellement sur les résultats d'essais cliniques, selon le principe de la médecine fondée sur les preuves. Des essais cliniques de qualité sont donc nécessaires afin de prodiguer les meilleurs soins possibles aux patients. Au cours des 50 dernières années, le nombre de publications scientifiques a fortement augmenté, principalement du fait de l'arrivée des chimiothérapies et des biothérapies. Parmi ces articles, on dénombre plus de 300 000 essais cliniques randomisés contrôlés entre 1969 et 2019 (Fig. 1).

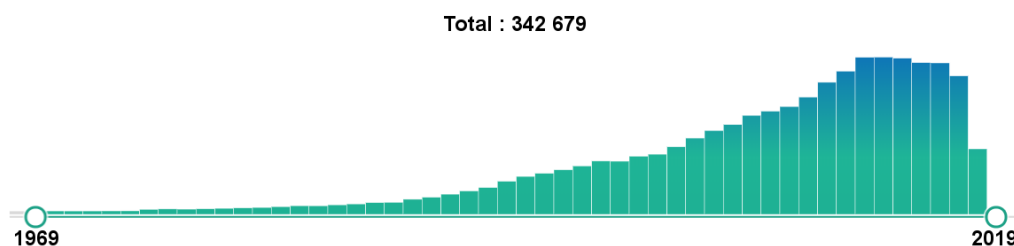


Figure 1 : Représentation annuelle du nombre d'essais cliniques randomisés contrôlés publiés dans le monde entre 1969 et 2019. Source : PubMed.

La qualité méthodologique des études reste extrêmement variable, même au sein de journaux internationaux avec comité de relecture. Les résultats d'un article sont basés sur les analyses statistiques effectuées, qui doivent être en accord avec le plan d'analyse statistique. Celui-ci doit être établi en amont de l'étude. Il précise le nombre de sujets à inclure dans l'étude et les analyses statistiques qui vont être effectuées, en fonction du type de données étudiées. Le calcul du nombre de sujets nécessaires (NSN) est donc une étape importante de la planification d'une étude. Le NSN doit être suffisamment important afin d'obtenir la puissance nécessaire pour

répondre à l'objectif de l'étude, mais ne doit pas inclure trop de sujets pour des raisons économiques et éthiques. En effet, inclure un nombre trop important de participants risque d'exposer inutilement certains sujets à de potentiels effets indésirables. Une surestimation de la taille de l'échantillon entraîne également un surcoût en allongeant la durée de l'étude.

1.2. Calcul du nombre de sujets nécessaires

L'objectif d'un essai clinique randomisé contrôlé est de mettre en évidence une différence de valeur du Critère de Jugement Principal (CJP) entre deux groupes (ou plus). L'analyse de cette différence se fait via des tests statistiques, selon une puissance (probabilité de mettre en évidence une différence qui existe réellement) et un risque de première espèce (probabilité de conclure à une différence qui n'existe pas réellement) donnés. Le calcul du NSN doit donc intégrer la puissance et le risque de première espèce (risque alpha). Il prend également en compte la latéralité du test statistique, c'est-à-dire l'expression de l'hypothèse alternative de manière bilatérale (valeur (A) \neq valeur (B)) ou unilatérale (valeur (A) < valeur (B) ou valeur (A) > valeur (B)). Tous ces paramètres sont fixés a priori et sont indépendants du type de variable étudiée.

Pour finir, le calcul du NSN doit également tenir compte de la distribution du CJP. Celle-ci peut en effet avoir une distribution normale (gaussienne), binomiale, exponentielle, de Poisson, ou encore géométrique. Pendant longtemps, la distribution de la variable étudiée n'a pas été prise en considération dans le calcul du NSN car cela nécessite souvent des calculs complexes. Néanmoins, les progrès de l'informatique dans le domaine des statistiques permettent désormais d'effectuer plus aisément ces calculs.

1.3. Généralités sur la Sclérose En Plaques

La Sclérose En Plaques (SEP) est une pathologie auto-immune chronique du système nerveux central. Elle se caractérise par une atteinte à la fois inflammatoire et démyélinisante, et une atteinte neurodégénérative. En France, on estime qu'environ 100 000 patients sont atteints de SEP (1). Elle touche en majorité les personnes jeunes (30 ans d'âge moyen au diagnostic), et préférentiellement les femmes (sex ratio à 2.5). On distingue plusieurs formes clinique de SEP, selon la présentation de la maladie (2). La forme récurrente-rémittente (RR) est la plus fréquente au diagnostic (80-85%). La maladie se caractérise par une succession d'épisodes de déficit neurologique appelés poussées, alternant avec des phases de rémission. Les symptômes régressent la plupart du temps, mais peuvent parfois perdurer, laissant une séquelle neurologique, source de handicap. En l'absence de traitement, la SEP-RR évolue vers une forme secondairement progressive (SP) dans au moins 50% des cas, après un délai plus ou moins long selon les patients. On distingue également les formes progressives primaires (PP) au cours desquelles le handicap s'installe progressivement dès le début de la maladie, sans poussées (10-15% des SEP).

La poussée est définie comme un épisode clinique monophasique, caractérisé par l'apparition aiguë ou subaiguë de symptôme(s) neurologique(s), en dehors de toute fièvre ou infection, perdurant au moins 24h et séparé au minimum de 30 jours d'un autre épisode neurologique (3). Dans les SEP de forme RR, le Taux Annualisé de Poussées (TAP) est un marqueur clinique d'activité de la maladie facile à calculer, et employé dans de nombreuses études. Il se définit comme le nombre de poussées observées au cours d'un temps d'observation donné, rapporté sur cette même période. Il existe une variation de la valeur du TAP inter-individuelle, mais également chez un même individu, le TAP diminuant en général après la ou les poussées conduisant au diagnostic (phénomène de régression à la moyenne) (4).

1.4. Calcul du nombre de sujets nécessaires basé sur le taux annualisé de poussées

Au sein d'une population, la distribution du TAP n'est pas gaussienne. En effet, il existe un nombre très important de patients n'ayant pas présenté de poussées pendant l'étude, ou n'en ayant présenté qu'une seule. Peu de patients ont en réalité une maladie très active se traduisant par des TAP élevés. Cette répartition du TAP se retrouve à la fois dans les essais cliniques (Annexe 2) (5–10), et chez les patients en vie réelle (Figure 2).

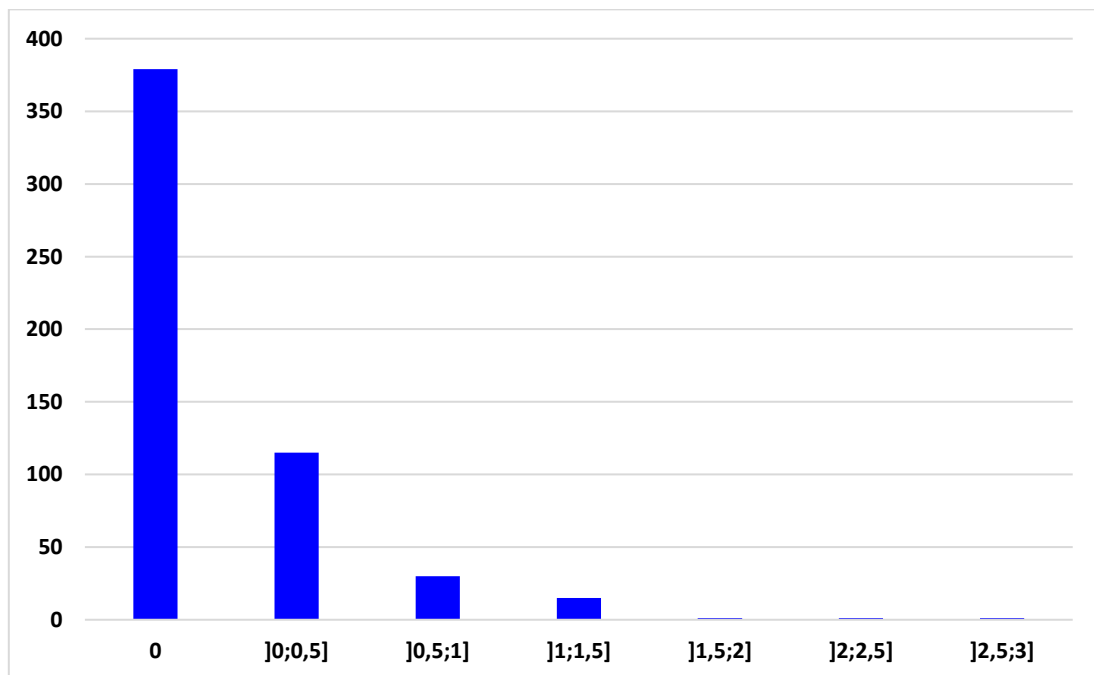


Figure 2 : Répartition du Taux Annualisé de Poussées chez les patients présentant une SEP-RR. Données extraites de la base de données locale du CHU de Clermont-Ferrand du registre européen EDMUS, entre septembre 2017 et septembre 2019, pour les patients traités ou non.

La loi de Poisson et la loi binomiale négative ont été proposées comme modèle statistique du TAP (11). La loi de Poisson ou « loi des événements rares », est définie par un paramètre λ correspondant à la fois à la moyenne et la variance. Le modèle binomial négatif est une généralisation de la loi de Poisson, qui prend en compte la surdispersion des données (12). Il permet d'analyser des jeux de données où la variabilité est plus grande que la moyenne de

l'échantillon. Le modèle binomial négatif semble correspondre au mieux à la distribution du TAP. Concernant les données présentées dans la figure 2, le TAP moyen est en effet de 0.2, avec un écart-type plus élevé, d'une valeur de 0.4. A titre de comparaison, la modélisation d'une distribution normale de paramètres (moyenne : 0.2 ; écart-type : 0.4) donnerait des valeurs négative de TAP (Figure 3).

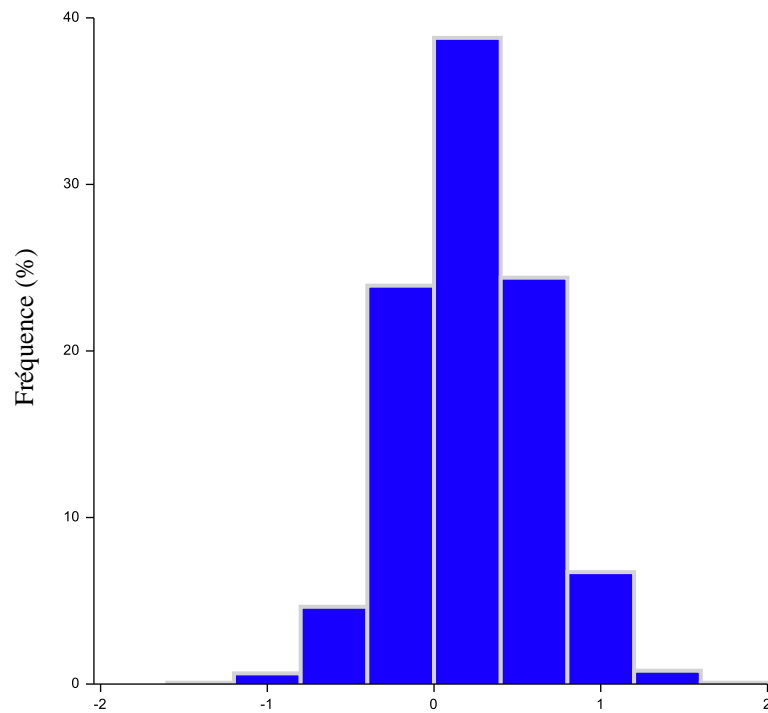


Figure 3. Simulation de données selon une loi normale.
Moyenne = 0.2 ; Ecart-type = 0.4

2. Hypothèses de travail et objectifs

La distribution du TAP selon un modèle binomial négatif nécessite l'utilisation de tests statistiques appropriés. Le calcul du NSN doit lui aussi être adapté à cette distribution. Or, les techniques de calcul basées sur un CJP suivant un modèle binomial négatif sont disponibles depuis plusieurs années (13), rendant possible son utilisation dans les essais cliniques ayant le TAP comme CJP.

Notre objectif principal est de déterminer si les méthodes utilisées pour le calcul du NSN et pour l'analyse du TAP sont adaptées à la distribution de ce dernier.

En cas d'inadéquation des techniques de calcul, notre objectif secondaire est de proposer une méthode de calcul prenant en compte la distribution du TAP.

3. Références bibliographiques

1. Foulon S, Weill A, Maura G, Dalichampt M, Debouverie M, Moreau T. Prévalence de la sclérose en plaques en France en 2012 et mortalité associée en 2013 à partir des données du Sniiram-PMSI. *Rev DÉpidémiologie Santé Publique*. 1 mars 2015;63:S17-8.
2. Lublin FD, Reingold SC. Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. *Neurology*. avr 1996;46(4):907-11.
3. Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol*. 2018;17(2):162-73.
4. Nicholas R, Straube S, Schmidli H, Pfeiffer S, Friede T. Time-patterns of annualized relapse rates in randomized placebo-controlled clinical trials in relapsing multiple sclerosis: a systematic review and meta-analysis. *Mult Scler Houndmills Basingstoke Engl*. sept 2012;18(9):1290-6.
5. Sorensen PS, Mellgren SI, Svenningsson A, Elovaara I, Frederiksen JL, Beiske AG, et al. NORDic trial of oral Methylprednisolone as add-on therapy to Interferon beta-1a for treatment of relapsing-remitting Multiple Sclerosis (NORMIMS study): a randomised, placebo-controlled trial. *Lancet Neurol*. juin 2009;8(6):519-29.
6. Giovannoni G, Comi G, Cook S, Rammohan K, Rieckmann P, Soelberg Sørensen P, et al. A placebo-controlled trial of oral cladribine for relapsing multiple sclerosis. *N Engl J Med*. 4 févr 2010;362(5):416-26.
7. Cohen JA, Barkhof F, Comi G, Hartung H-P, Khatri BO, Montalban X, et al. Oral fingolimod or intramuscular interferon for relapsing multiple sclerosis. *N Engl J Med*. 4 févr 2010;362(5):402-15.
8. O'Connor P, Wolinsky JS, Confavreux C, Comi G, Kappos L, Olsson TP, et al. Randomized trial of oral teriflunomide for relapsing multiple sclerosis. *N Engl J Med*. 6 oct 2011;365(14):1293-303.
9. Sorensen PS, Lycke J, Erälinna J-P, Edland A, Wu X, Frederiksen JL, et al. Simvastatin as add-on therapy to interferon β -1a for relapsing-remitting multiple sclerosis (SIMCOMBIN study): a placebo-controlled randomised phase 4 trial. *Lancet Neurol*. août 2011;10(8):691-701.
10. Gold R, Giovannoni G, Selmaj K, Havrdova E, Montalban X, Radue E-W, et al. Daclizumab high-yield process in relapsing-remitting multiple sclerosis (SELECT): a randomised, double-blind, placebo-controlled trial. *Lancet Lond Engl*. 22 juin 2013;381(9884):2167-75.
11. Wang YC, Meyerson L, Tang YQ, Qian N. Statistical methods for the analysis of relapse data in MS clinical trials. *J Neurol Sci*. 15 oct 2009;285(1-2):206-11.
12. Jourdan A, Kokonendji CC. Surdispersion et modèle binomial négatif généralisé. *Rev Stat Appliquée*. 2002;50(3):73-86.
13. Zhu H, Lakkis H. Sample size calculation for comparing two negative binomial rates. *Stat Med*. 10 févr 2014;33(3):376-87.

ARTICLE

Estimation of sample size in Multiple Sclerosis Randomized Controlled Trials studying Annualized Relapse Rate: a systematic review

L. Poncet-Megemont¹, B. Pereira², P. Clavelou¹⁻³, and X. Moisset¹⁻³

¹Service de Neurologie, CHU de Clermont-Ferrand, Université Clermont Auvergne, F-63000, Clermont-Ferrand, France

²Biostatistics Unit, CHU Clermont-Ferrand, Délégation Recherche Clinique and Innovation, Clermont-Ferrand, France.

³Université Clermont Auvergne, Inserm Neuro-Dol U1107, F-63000 Clermont-Ferrand, France

Glossary :

ARR : Annualised Relapse Rate ; **EDSS** : Expanded Disability Status Scale ; **MS** : Multiple Sclerosis ; **NB** : Negative Binomial ; **RCT** : Randomized Controlled Trial ; **RRMS** : Relapsing Remitting MS ; **SD** : Standard Deviation ; **SS** : Sample Size ; **SSE** : Sample Size Estimation.

Abstract

Background : In randomized controlled trials (RCTs), the sample size estimation (SSE) is an important step of the planning of statistical analyses. The calculation method must be adapted to the distribution of the studied variable. In multiple sclerosis (MS) studies, the Annualised Relapse Rate (ARR) is a marker of the disease activity, which distribution follow a negative binomial (NB) distribution. Our objectives were to determine whether the SSE and the analysis of ARR in phase 3 and 4 RCTs of drug therapies for relapsing remitting MS were consistent with the distribution of the ARR, and to compare the SSE across several statistical distributions.

Methods : A systematic review of the literature included phase 3 and 4 RCTs which primary endpoint was ARR in relapsing remitting MS, published since 2008 in scientific journals with an Impact Factor ≥ 4 . We checked whether the SSE and the ARR analyses were congruent with a NB model. We also performed SSE based on normal and NB distribution, using the data collected from the studies.

Results : Nineteen articles corresponding to 20 studies were selected. Nine of them indicated the method used for SSE. Of these, only 4 used a model derived from a Poisson or NB distribution. In contrast, the ARR analyses were based on these two models in 19 out of 20 studies. SSE using identical parameters showed highly variable results, depending on the ARR and dispersion used.

Conclusions : Our study pointed out that the SSE in MS RCTs studying ARR is rarely performed with the accurate model. The complexity of the NB model creates variable difference compared to the gaussian model. As it is difficult to predict these differences, the use of a NB model seems mandatory to optimize the number of included patients.

Introduction

New drugs are constantly in development, resulting in a large number of clinical studies, which are often long and expensive. In the last two decades, a growing number of methodological papers raised awareness of issues and gaps in the scientific publications' methodology, mostly in Randomized Controlled Trials (RCTs). Transparency and public access to study protocol seems to be the first step for higher quality of published data (1).

During the planning of a new study, the Sample Size Estimation (SSE) is a key step, on which the reliability of the results depends. In Multiple Sclerosis (MS) studies, many clinical parameters can be studied, depending on the course of the disease. In Relapsing Remitting MS (RRMS), Annualized Relapse Rate (ARR) is often chosen as primary outcome in phase 3 and phase 4 RCTs. The number of relapses directly impacts the quality of life of the patient, which explain why it is clinically relevant. Moreover, ARR is easily collected and has the advantage of quickly showing a difference between treatment groups resulting in shorter studies compared to outcomes based on disability. Moreover, a new relapse is the opportunity to reassess the treatment of a patient and is often a chance for choosing a more effective immunotherapy.

ARR is a parameter which is non-normally distributed. In fact, due to the small number of relapses per patient each year, ARR follows a distribution that obeys to the law of rare events. It can be modeled through Poisson derived distributions. A Poisson distribution is a discrete probability distribution defined with a single parameter λ which is both the mean and the variance, which results in an absence of data dispersion. Moreover, in a Poisson model each event is independent from others. The quasi-likelihood Poisson distribution and the negative binomial regression are Poisson derived distribution which consider overdispersion of data, allowing the variance to be greater than the mean. Overdispersion results from heterogeneity of relapse rate between patient, and from contagion (an early relapse involves a higher risk of

subsequent ones), both mechanisms leading to a statistical link between relapses (2). The quasi-likelihood Poisson distribution and the negative binomial model are both suitable to study ARR, and it is currently recommended to use negative binomial regression for the design and the analysis of relapse rate (2).

SSE should therefore be based on a negative binomial regression, and not on other distributions classically used, such as a gaussian distribution, which do not correspond to the observed distribution of ARR.

Based on this observation, we conducted a systematic review to investigate if the statistical analysis plan (the SSE) and the statistical analysis of RCTs using ARR as their primary outcome was congruent with the ARR distribution. The secondary outcome was to propose an appropriate SSE calculation and check for any difference with the real inclusions performed.

Methods

Reviewing of studies and data collection

This systematic review was made using PubMed to identify the clinical trials of interest. We chose to focus on general and specialized journals with high visibility and large readership, with an Impact Factor superior or equal to 4 at the planning of this study (2019). In these journals, we only reviewed article from the last decade, precisely from January 1st, 2008 to December 31st, 2019. We then selected only the phase 3 and 4 randomised controlled trials dealing with RRMS in adults, and which primary endpoint was the ARR. If there was no proper estimation of the sample size (SS), the study was excluded. Moreover, if ARR was a co-primary endpoint with another outcome, the clinical trial could be included if the SS was calculated

based on ARR data and not on the data of the co-primary endpoint. Otherwise, the paper was not included.

In the selected studies, we extracted data from the original article. When information was missing, we looked for it in every document available online : supplementary appendix, study protocol, data from websites (*clinicaltrial.gov* or *trialsjournal.biomedcentral.com*). If still missing, we tried to recover the expected standard deviation for each group by sending mails to the corresponding authors and the last authors.

Endpoints of the study

Our primary endpoint was to gather the statistical methods used for the SSE and the statistical analysis of each selected RCT, to check if it corresponds well to the ARR distribution. Our secondary endpoint was to reassess the SS for each study, using two different methods and to compare our results to the estimated SS published in the studies. The first SSE was based on a normal distribution of the ARR, and the second one relies on a negative binomial distribution. For each method, we performed a SSE based on the data provided in the statistical plan analysis, and another SSE based on the data really observed in the study. Exploratory endpoints were also collected to have an overview of the last decade RCTs in MS : power, type I error (alpha risk), number of tails, expected ARR in the control group and in the intervention group, standard deviation of the ARR in both groups, and expected dropout rate.

Sample size estimations

When possible, we performed the SSE based on expected ARR and expected dispersion specified in the statistical analysis plan, relying on a normal distribution (Method I). Next, for each included study, we performed SSE based on the distribution observed in the studies,

obtaining the standard deviation (SD) and the negative binomial dispersion parameter k from the 95% confidence interval of the ARR. We calculated the SS based on the expected ARR (Method II and III, respectively for normal and NB distribution) and the observed ARR (Method IV and V, respectively for normal and NB distribution). For all SSE, we kept type 1 error, power and group allocation indicated by the authors. Finally, we simulated SSE based on NB regression to study the factors influencing SS. The SSE relying on a normal distribution were performed with G*Power software (version 3.1.9.4). The SSE based on a negative binomial distribution were assessed using Stata software (version 13.1, StataCorp, College Station, TX), thanks to the package *power tworates_zhu* provided by Mark Chatfield (available online), derived from the work of Zhu and Lakkis (3). SSE simulations were achieved with PASS 2020 (restricted free trial, NCSS). This review is registered with PROSPERO, under acceptance.

Results

Studies selected

Of the 4474 articles identified, only 942 were published in pre-selected journals. Finally, 19 articles were included in this systematic review (Figure 1), corresponding to 20 studies as Hauser *et al.* work presented the results of two independent studies, namely OPERA I and II (4).

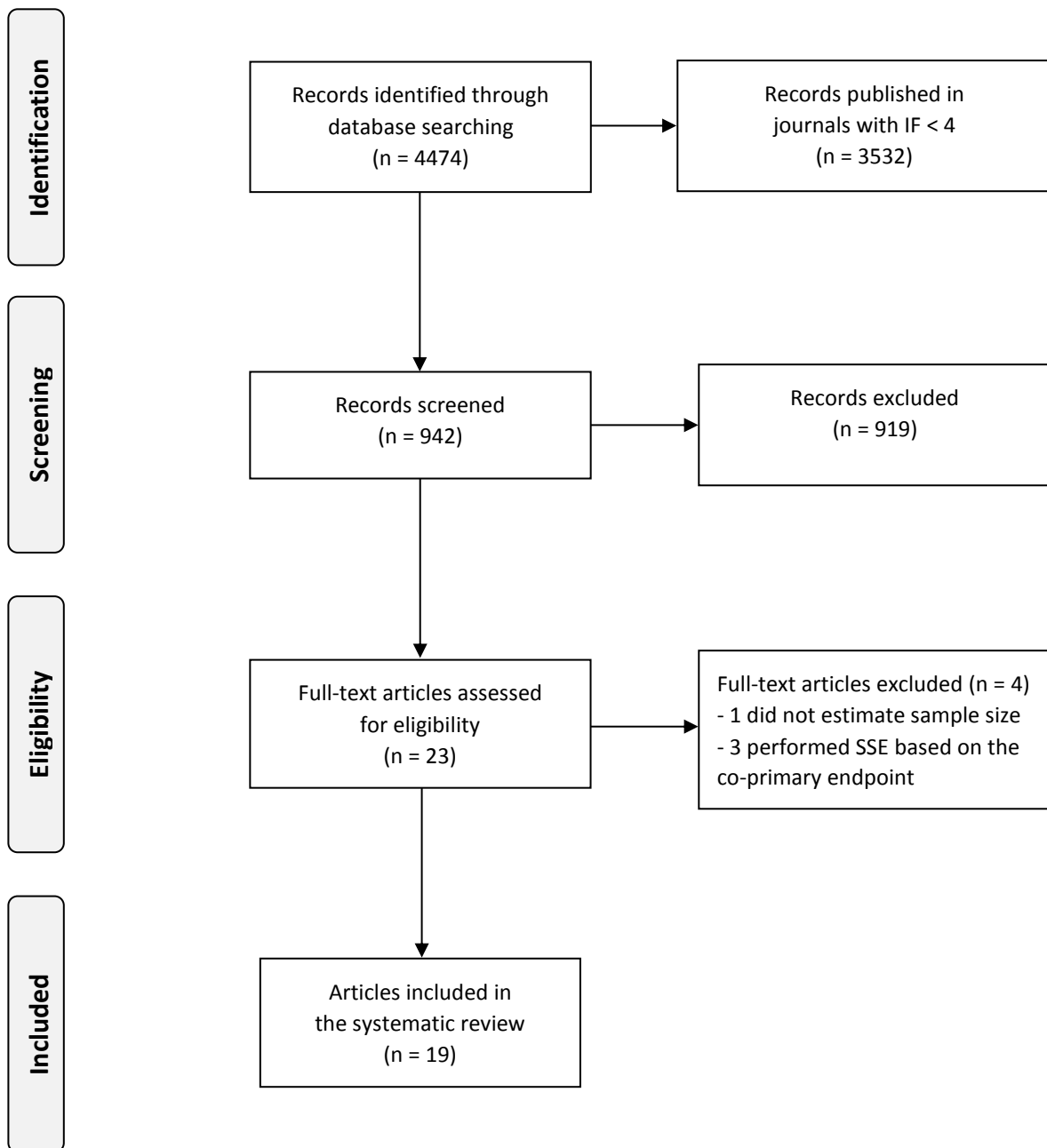


Figure 1. Flow diagram for inclusion of studies in the systematic review.

Statistical method used for SSE and analysis or ARR

Within the “Method” part, the exhaustivity of the statistical analysis section was quite heterogeneous between studies, depending on the journal and the year of publication. This statement also applied to the availability of online supplementary material. The descriptions of the method used tend to improve across time. Only 9 out of 20 studies indicated the test used to estimate their SS. Four of them used a quasi-likelihood Poisson distribution, or a negative

binomial distribution model (5–8). Four indicated the use of a Student t-test (4,9,10), and one of a Wilcoxon-Mann-Whitney rank-sum test (11). For the other eleven, the statistical method used for SSE was not specified.

Journal	Year of pub.	Authors	Reference
<i>Annals of Neurology</i>	2011	G. Comi et al.	(12)
	2013	F.D. Lublin et al.	(9)
	2013	O. Khan et al.	(13)
<i>Lancet</i>	2013	R. Gold et al.	(5)
<i>Lancet Neurology</i>	2009	P.S. Sorensen et al.	(14)
	2011	P.S. Sorensen et al.	(15)
	2014	C. Confavreux et al.	(6)
	2014	P.A. Calabresi et al.	(16)
	2014	P.A. Calabresi et al.	(17)
	2016	R. R. Voskuhl et al.	(18)
<i>Multiple Sclerosis Journal</i>	2009	E. Havrdova et al.	(19)
<i>New England Journal of Medicine</i>	2010	L. Kappos et al.	(20)
	2010	J.A. Cohen et al.	(11)
	2010	G. Giovannoni et al.	(10)
	2011	P. O'Connor et al.	(21)
	2012	G. Comi et al.	(7)
	2012	R.J. Fox et al.	(22)
	2015	L. Kappos et al.	(8)
	2017	S.L. Hauser et al.	(4)

Table 1. Publication characteristics of the articles included in the systematic review.

On the contrary, the methods used to analyse the ARR were quite homogeneous and congruent with the statistical distribution of ARR (see Table 2). Indeed, a quarter of the ARR analyses were based on a Poisson model, and more than half relied on a quasi-likelihood Poisson distribution or a negative binomial regression. Some authors indicated in the online study protocol their will to analyse ARR with one of these last two methods, and to change for the other one depending on the dispersion of data (7,22). The most common covariates were age at

baseline, EDSS (Expanded Disability Status Scale), number of relapses during a period preceding inclusion, country or geographic region, and study group (when accurate).

Year of pub. Authors	Statistical test for SSE	Statistical test for analysis
2009 E. Havrdova et al.	NA	Poisson regression
2009 P.S. Sorensen et al.	NA	Poisson regression
2010 L. Kappos et al.	NA	Negative binomial regression
2010 J.A. Cohen et al.	Wilcoxon-Mann-Whitney rank-sum test	Negative binomial regression
2010 G. Giovannoni et al.	Student t-test	Poisson regression
2011 G. Comi et al.	NA	Quasi-likelihood (overdispersed) Poisson regression
2011 P. O'Connor et al.	NA	Poisson regression
2011 P.S. Sorensen et al.	NA	Quasi-likelihood (overdispersed) Poisson regression
2012 G. Comi et al.	Quasi-likelihood (overdispersed) Poisson regression model	Quasi-likelihood (overdispersed) Poisson regression
2012 R.J. Fox et al.	NA	Negative binomial regression
2013 R. Gold et al.	Negative binomial regression model	Negative binomial regression
2013 F.D. Lublin et al.	Student t-test	Cox proportional hazards
2013 O. Khan et al.	NA	Quasilielihood (overdispersed) Poisson regression
2014 P.A. Calabresi et al.	NA	Negative binomial regression
2014 P.A. Calabresi et al.	NA	Negative binomial regression
2014 C. Confavreux et al.	Quasi-likelihood (overdispersed) Poisson regression model	Poisson regression
2015 L. Kappos et al.	Negative binomial regression model	Negative binomial regression
2016 R. R. Voskuhl et al.	NA	Negative binomial regression
2017 S.L. Hauser et al.	Student t-test	Negative binomial regression

Table 2. Statistical models used for SSE and for analysis of the ARR.
NA : Not Available

Comparison of sample size estimation

The planned SS, the number of subjects included, and the SSE assessing normal and negative binomial distribution are detailed in *Supplementary Data*. For the 9 studies which indicated the expected SD, the planned SS was superior to the SS based on a normal distribution using expected ARR and SD in 5 studies (18.7% average increase), and inferior in 4 studies (19.6% average decrease). Method IV required more subjects than method II in 11 out of 16 studies, the only parameter changing being the ARR of both groups, resulting in a smaller between ARR difference for these 11 studies. When based on expected ARR, the NB model required more participants than the gaussian model in 75% of the studies (method II vs method III).

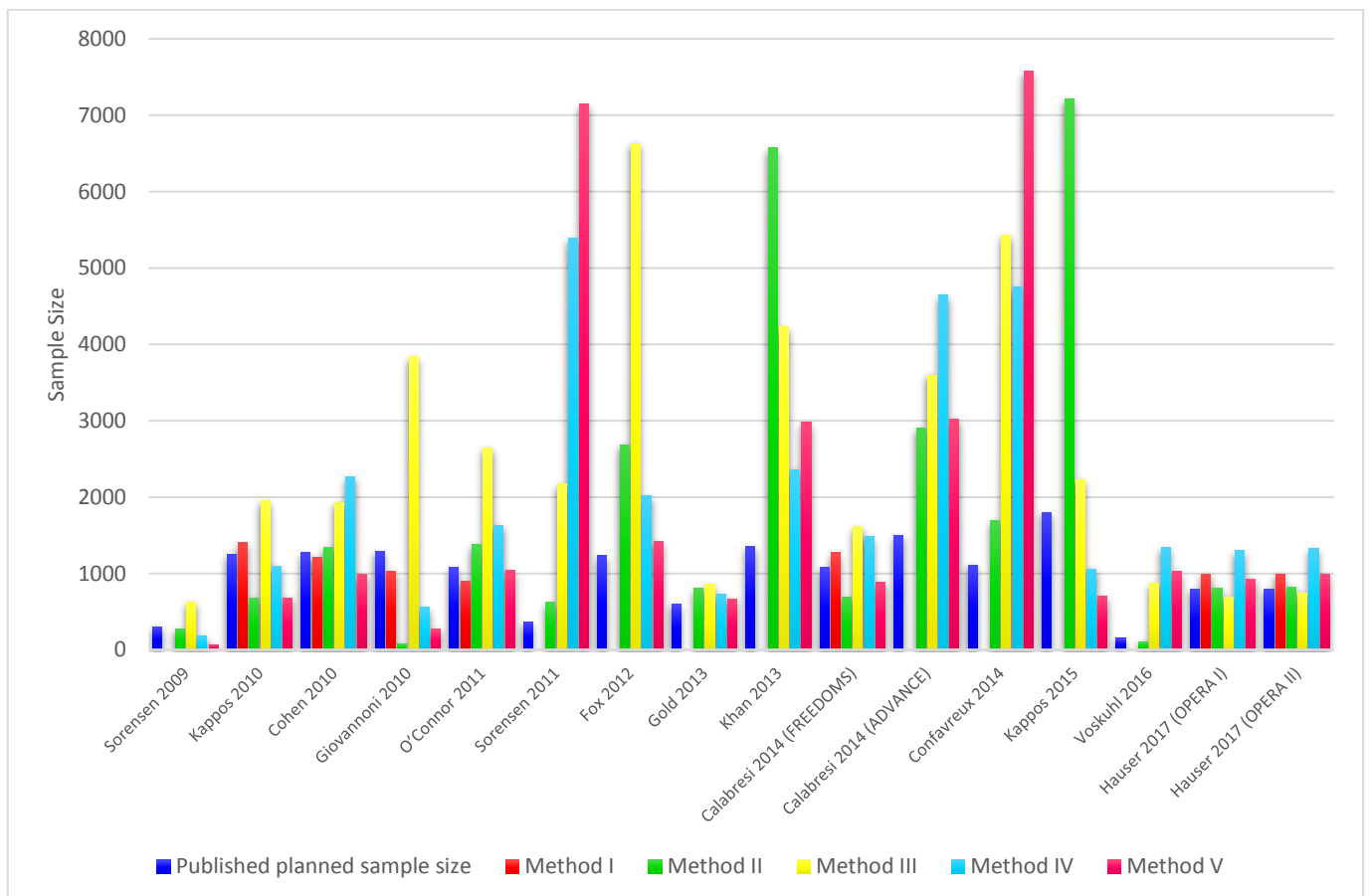


Figure 2. Comparison of SSE with the different methods.

Method I : Estimation with a normal distribution based on expected parameters and expected dispersion ; *Method II* : Estimation with a normal distribution based on expected parameters and observed dispersion ; *Method III* : Estimation with a negative binomial distribution based on expected parameters and observed dispersion ; *Method IV* : Estimation with a normal distribution based on observed parameters ; *Method V* : Estimation with a negative binomial distribution based on observed parameters. Studies that did not indicate the dispersion of the ARR observed (CI 95% or else) could not be computed. Published planned sample size are given for information and not for comparison purpose.

On the contrary, when based on real ARR observed in the studies, the NB model required less participants than the gaussian model (method IV vs method V), only 3 studies showing opposite results (6,13,15). Taking these results together, the SSE with normal and NB models were divergent for 6 studies (Figure 2). when the dispersion and the ARR of both control and treatment group changed (increase of SS with the normal model, and decrease with the NB model) (10,11,16,17,20,21).

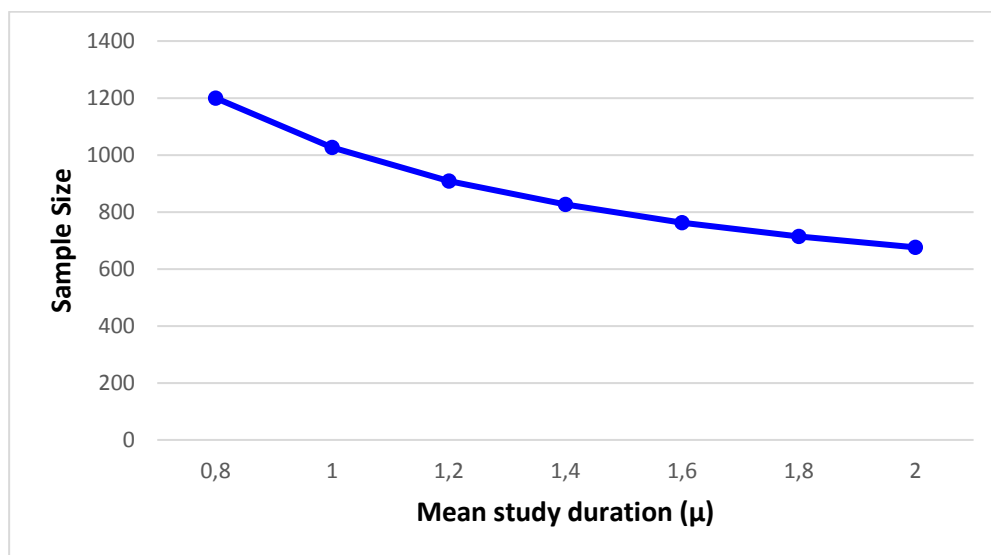


Figure 3. SS variation according to the length of the study.

All SSE are performed with two tails, a type 1 error of 0.15, a power of 96,5%, and a NB dispersion parameter $k=2$.

The SSE formula following a NB distribution is complex and integrates parameters that are not considered in SSE based on normal model, such as study duration. We simulated SSE with fixed parameters, studying different duration. It shows that the needed SS is greater for short studies than for longer ones (Figure 3).

We then performed multiple SSE by varying the other parameters that are part of the negative binomial SSE formula. These simulations indicate that the SS increase when the ARR of the

control group is lower, when the risk ratio increases (low between-group ARR difference), and when data dispersion increases (Figure 4).

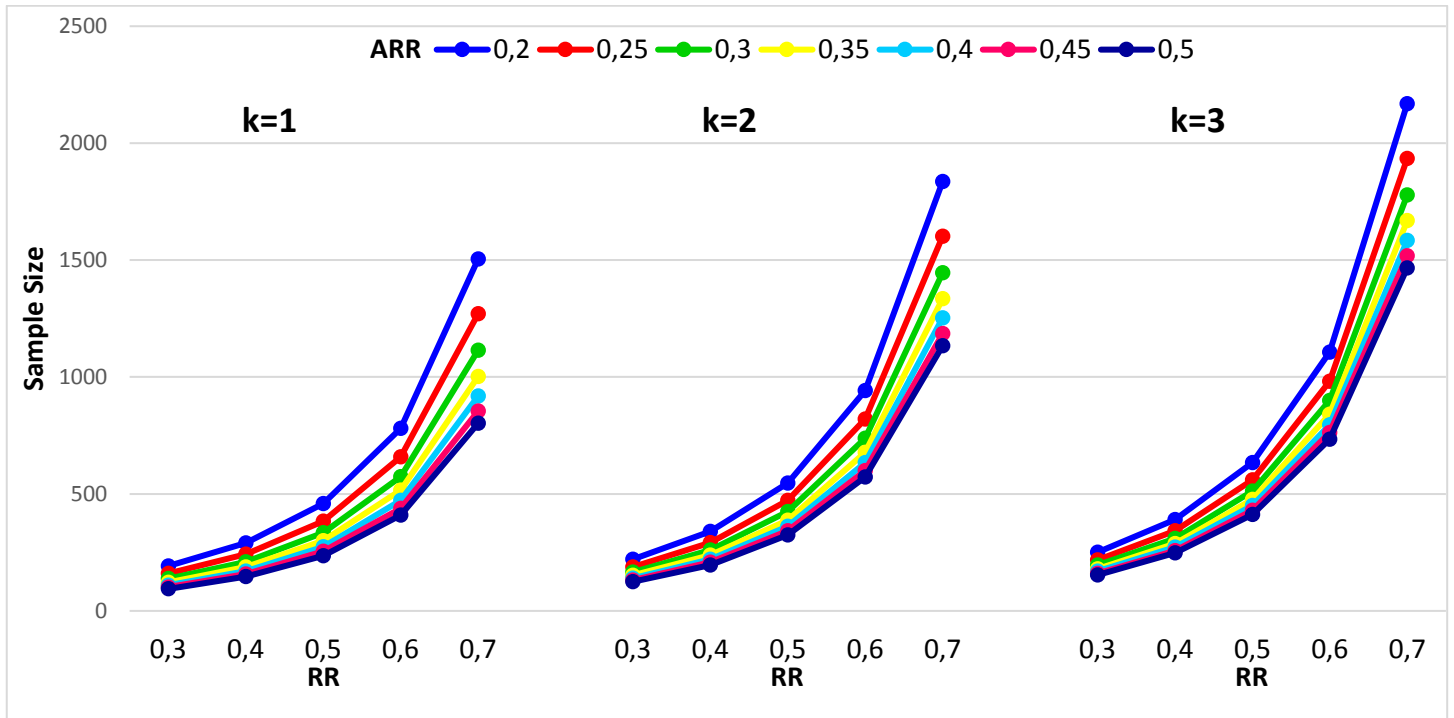


Figure 4. Simulation of SSE following a negative binomial distribution, with different basal ARR (control group), and different risk ratio, based on 3 values of k

All SSE are performed with two tails, a type 1 error of 0.15, a power of 96,5%, and a mean duration of the study of 1.7 years. ARR : Annualized Relapse Rate. k : negative binomial dispersion parameter. RR : Risk Ratio

Other characteristics of the studies

As exploratory endpoints, we were also interested in other statistical parameters that are needed for SSE (see *Supplementary Data*). Here again, results were very heterogeneous. The value of power was widely variable, while alpha risk was usually set at 5%. In studies with ≥ 3 treatment groups, more than 1/3 did not performed or did not indicated how they managed alpha risk inflation due to multiple group comparisons. The expected ARR values of the control group were very different among the studies, ranging from 0.35 to 2.1 for placebo and from 0.27 to 1.1 in interferon groups. When indicated, dispersion of ARR was also highly variable, with SD

ranging from 0.6 to 2.02 (median 0.87), whereas in the results, it was more constant (0.525 to 1.48 ; median 0.776). The observed SD was superior to the expected SD in only 2 studies among the 7 for which the comparison could be performed. Among the 8 articles in which the dispersion parameter was clearly identified, only one used a different value for each treatment group, to reflect the drug effectiveness. Other data such as number of tails, expected ARR reduction and dropout rate are listed in *Supplementary Data*. The strategy used for the management of missing data was clearly indicated in most of the studies.

Discussion

General findings

This systematic review is the first one focusing on the methodology of SSE in MS studies taking ARR as primary outcome. The distribution of the variable of interest is of major importance in the analysis of a result. In trials studying ARR in MS, this fact is usually properly considered. Nonetheless, we show for the first time that SSE is infrequently performed with the most appropriate statistical methods, even in well-conducted studies analysed properly and published in major medical journals, although this step is also of the greatest importance in the study design. Thus, should we stop reading articles when the SSE is not performed with the most appropriate method?

We also report a lack of details in the statistical methods of RCTs. It is therefore difficult to know which methods and which parameters were used, and to check the quality of the analyses.

SSE and determining parameters

In our review, there was an important difference between expected and observed ARR (up to 84%, mean : 46%), the observed ARR being lower in most of the studies (18/20). Moreover, we found that the observed ARR was extremely variable among studies. This variability is present between studies and across time (23,24), and explains the impossibility to set a standardized drug specific value of ARR that could be used for future studies. Classically, trialists use the ARR obtained in the phase 2 study of the tested drug (and extension studies if available), which enrol patient with very homogenous demographic characteristics (11,16,20). Unfortunately, when based on a phase 2 study, the expected ARR in the treatment group is often higher than the ARR finally observed in the phase 3 study, and the dispersion is lower, resulting in biased SSE. Indeed, phase 2 studies are often imaging studies lasting one year, while RCTs based on clinical parameters last longer (up to 3 years). In the first 3 to 6 months after treatment initiation, the disease may still be active due to delayed treatment effect, and moreover MS is more active during the first year of RCTs than during the second year (25) due to the regression to the mean phenomena.

To improve the uncertainty about the value of ARR during the planning of future studies, there is a growing interest for SSE using a network meta-analysis based on past trials of different drugs with the same primary outcome (26). This method seems to need less subjects per arm (27) and can be interesting to determine expected ARR for the testing of drugs in future trials. To go further, some authors proposed to include data from RCTs and from real world data (i.e. data from observational and other interventional studies) in a Bayesian network meta-analysis to simulate SSE (28). To our opinion, the mix of real-world data and RCTs data produce a study population with different clinical and disease characteristics and should not be used for the SSE of a RCT with strict inclusion criteria and treatment attribution based only on randomization.

Comparison of two ratio with a negative binomial model

Because they assume overdispersion of counted data, the quasi-likelihood Poisson model and the negative binomial regression fit perfectly the ARR distribution, the last one being preferred because the variance-to-mean ratio of ARR is not constant with the mean (2).

Our results show that SSE anticipating a NB distribution is highly variable because of the complexity of its formula, and do not necessarily follow the variations observed with a gaussian model. Indeed, parameters such as ARR in the control group, risk ratio, study duration and data dispersion influence the SSE. It explains why for 3 studies, the SS was greater with the NB model than with the normal distribution, probably due to multiple causes : short studies with wide data dispersion, and moreover one with high power (6) and another with extreme value or risk ratio (13).

The formula of SSE based on NB distribution is easily usable, and do not require powerful software. It integrates more parameters than SSE based on a normal distribution. For this, it is very complete, but needs very precise value to avoid producing aberrant SS. As detailed above, precise value of ARR and its dispersion can be easily obtained. Based on these results, we can suppose that changing the way SSE is calculated would allow to enrol a lower number of patients in phase 3 and phase 4 studies. Such reduction in the SS allow to minimize the risk of a study, exposing the lowest necessary number of patients, and minimizing the cost and the length of such studies.

Our study has limitations. First, there was a huge amount of missing data in lots of articles, even if we collected data from all available documents. Therefore, we were not able to properly perform SSE with the 5 methods. We join the current call for transparency, encouraging authors to register their study protocols online, and to integrate as much information on the results in supplementary materials. Secondly, due to the complexity of the SSE formula, it was difficult

to perform proper comparison of SSE based on normal and NB models, requiring us to estimate SS from the observed dispersion only. Although providing divergent results, we observed a clear decrease of the SS for most of the studies when based on observed parameters. The purpose of this work was not to question the results of the reviewed studies, and we did not performed comparison of our SSE with the initially planned SS, as it is not pertinent (29).

Alternative to ARR and proposal for future studies

As an alternative to ARR as primary outcome, Time To First Relapse (TTFR) or number of relapse-free patients (RFP) seems seducing for trialists because it may require less subjects to show a difference between treatment groups, without modifying the power of MS studies, particularly in studies in which ARR is low (30). However, such outcome rises the issue of the continuation of disease activity during the first months after the start of the study due to delayed treatment effect. Furthermore, in studies using TTFR or RFP, data are usually censored after the first relapse, as the patient can be switched to another active treatment. This may lead to missing data for other outcome collected at different milestone (safety, disability). For these reasons, clinical studies based on ARR, or integrating ARR in a composite score (such as NEDA : Non-Evidence of Disease Activity), should be preferred.

We have seen that performing SSE based on a NB distribution could allow pharmaceutical companies to enrol less patients. Another method frequently proposed is to design more adaptative trials to fit more precisely to the real SS needed. Intermediate analyses are adaptative methods but increase type I error, may induce stopping of the study for futility or with rejection of the null hypothesis. Blinded Sample Size Reestimation (BSSR) by an independent Data Monitoring Committee at midpoint of the study does not have this inconvenient as the reestimation is blind and no statistical analyses are performed (31). BSSR allow to adapt to the

ARR really measured in the study population and is applicable to MS (32). It has been previously used in ADVANCE study (17) and could have been used in studies showing a much greater reduction of ARR (>20%) that initially planned, which could have allow to reduce the SS ((8,10,14,20,22).

For MS RCTs studying ARR, SSE based on negative binomial distribution is recommended. This method is available, and our study highlights the importance to use the same method for both the planning of the study and the analyses of the results, in order to optimize the number of patients to be included and allow to present reliable data. These conclusions can be extended to other domains in MS studies, such as imaging studies in which the lesion count follows a negative binomial distribution (33). SSE based on a normal distribution should therefore be banned from statistical plan analysis of MS studies.

References

1. Chan A-W, Hróbjartsson A. Promoting public access to clinical trial protocols: challenges and recommendations. *Trials* [Internet]. 17 févr 2018 [cité 30 nov 2019];19. Disponible sur: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5816550/>
2. Wang YC, Meyerson L, Tang YQ, Qian N. Statistical methods for the analysis of relapse data in MS clinical trials. *J Neurol Sci*. 15 oct 2009;285(1-2):206-11.
3. Zhu H, Lakkis H. Sample size calculation for comparing two negative binomial rates. *Stat Med*. 10 févr 2014;33(3):376-87.
4. Hauser SL, Bar-Or A, Comi G, Giovannoni G, Hartung H-P, Hemmer B, et al. Ocrelizumab versus Interferon Beta-1a in Relapsing Multiple Sclerosis. *N Engl J Med*. 19 2017;376(3):221-34.
5. Gold R, Giovannoni G, Selmaj K, Havrdova E, Montalban X, Radue E-W, et al. Daclizumab high-yield process in relapsing-remitting multiple sclerosis (SELECT): a randomised, double-blind, placebo-controlled trial. *Lancet Lond Engl*. 22 juin 2013;381(9884):2167-75.
6. Confavreux C, O'Connor P, Comi G, Freedman MS, Miller AE, Olsson TP, et al. Oral teriflunomide for patients with relapsing multiple sclerosis (TOWER): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Neurol*. mars 2014;13(3):247-56.
7. Comi G, Jeffery D, Kappos L, Montalban X, Boyko A, Rocca MA, et al. Placebo-controlled trial of oral laquinimod for multiple sclerosis. *N Engl J Med*. 15 mars 2012;366(11):1000-9.
8. Kappos L, Wiendl H, Selmaj K, Arnold DL, Havrdova E, Boyko A, et al. Daclizumab HYP versus Interferon Beta-1a in Relapsing Multiple Sclerosis. *N Engl J Med*. 8 oct 2015;373(15):1418-28.
9. Lublin FD, Cofield SS, Cutter GR, Conwit R, Narayana PA, Nelson F, et al. Randomized study combining interferon and glatiramer acetate in multiple sclerosis. *Ann Neurol*. mars 2013;73(3):327-40.
10. Giovannoni G, Comi G, Cook S, Rammohan K, Rieckmann P, Soelberg Sørensen P, et al. A placebo-controlled trial of oral cladribine for relapsing multiple sclerosis. *N Engl J Med*. 4 févr 2010;362(5):416-26.
11. Cohen JA, Barkhof F, Comi G, Hartung H-P, Khatri BO, Montalban X, et al. Oral fingolimod or intramuscular interferon for relapsing multiple sclerosis. *N Engl J Med*. 4 févr 2010;362(5):402-15.
12. Comi G, Cohen JA, Arnold DL, Wynn D, Filippi M, FORTE Study Group. Phase III dose-comparison study of glatiramer acetate for multiple sclerosis. *Ann Neurol*. janv 2011;69(1):75-82.

13. Khan O, Rieckmann P, Boyko A, Selmaj K, Zivadinov R, GALA Study Group. Three times weekly glatiramer acetate in relapsing-remitting multiple sclerosis. *Ann Neurol.* juin 2013;73(6):705-13.
14. Sorensen PS, Mellgren SI, Svenningsson A, Elovaara I, Frederiksen JL, Beiske AG, et al. NORdic trial of oral Methylprednisolone as add-on therapy to Interferon beta-1a for treatment of relapsing-remitting Multiple Sclerosis (NORMIMS study): a randomised, placebo-controlled trial. *Lancet Neurol.* juin 2009;8(6):519-29.
15. Sorensen PS, Lycke J, Erälinna J-P, Edland A, Wu X, Frederiksen JL, et al. Simvastatin as add-on therapy to interferon β -1a for relapsing-remitting multiple sclerosis (SIMCOMBIN study): a placebo-controlled randomised phase 4 trial. *Lancet Neurol.* août 2011;10(8):691-701.
16. Calabresi PA, Radue E-W, Goodin D, Jeffery D, Rammohan KW, Reder AT, et al. Safety and efficacy of fingolimod in patients with relapsing-remitting multiple sclerosis (FREEDOMS II): a double-blind, randomised, placebo-controlled, phase 3 trial. *Lancet Neurol.* juin 2014;13(6):545-56.
17. Calabresi PA, Kieseier BC, Arnold DL, Balcer LJ, Boyko A, Pelletier J, et al. Pegylated interferon β -1a for relapsing-remitting multiple sclerosis (ADVANCE): a randomised, phase 3, double-blind study. *Lancet Neurol.* juill 2014;13(7):657-65.
18. Voskuhl RR, Wang H, Wu TCJ, Sicotte NL, Nakamura K, Kurth F, et al. Estriol combined with glatiramer acetate for women with relapsing-remitting multiple sclerosis: a randomised, placebo-controlled, phase 2 trial. *Lancet Neurol.* janv 2016;15(1):35-46.
19. Havrdova E, Zivadinov R, Krasensky J, Dwyer MG, Novakova I, Dolezal O, et al. Randomized study of interferon beta-1a, low-dose azathioprine, and low-dose corticosteroids in multiple sclerosis. *Mult Scler Houndmills Basingstoke Engl.* août 2009;15(8):965-76.
20. Kappos L, Radue E-W, O'Connor P, Polman C, Hohlfeld R, Calabresi P, et al. A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *N Engl J Med.* 4 févr 2010;362(5):387-401.
21. O'Connor P, Wolinsky JS, Confavreux C, Comi G, Kappos L, Olsson TP, et al. Randomized trial of oral teriflunomide for relapsing multiple sclerosis. *N Engl J Med.* 6 oct 2011;365(14):1293-303.
22. Fox RJ, Miller DH, Phillips JT, Hutchinson M, Havrdova E, Kita M, et al. Placebo-controlled phase 3 study of oral BG-12 or glatiramer in multiple sclerosis. *N Engl J Med.* 20 sept 2012;367(12):1087-97.
23. Inusah S, Sormani MP, Cofield SS, Aban IB, Musani SK, Srinivasasainagendra V, et al. Assessing changes in relapse rates in multiple sclerosis. *Mult Scler Houndmills Basingstoke Engl.* déc 2010;16(12):1414-21.
24. Steinworth SM, Röver C, Schneider S, Nicholas R, Straube S, Friede T. Explaining temporal trends in annualised relapse rates in placebo groups of randomised controlled

- trials in relapsing multiple sclerosis: systematic review and meta-regression. *Mult Scler Houndmills Basingstoke Engl.* oct 2013;19(12):1580-6.
25. Nicholas R, Straube S, Schmidli H, Pfeiffer S, Friede T. Time-patterns of annualized relapse rates in randomized placebo-controlled clinical trials in relapsing multiple sclerosis: a systematic review and meta-analysis. *Mult Scler Houndmills Basingstoke Engl.* sept 2012;18(9):1290-6.
 26. Tramacere I, Del Giovane C, Salanti G, D'Amico R, Filippini G. Immunomodulators and immunosuppressants for relapsing-remitting multiple sclerosis: a network meta-analysis. *Cochrane Database Syst Rev.* 18 sept 2015;(9):CD011381.
 27. Salanti G, Nikolakopoulou A, Sutton AJ, Reichenbach S, Trelle S, Naci H, et al. Planning a future randomized clinical trial based on a network of relevant past trials. *Trials.* 11 juill 2018;19(1):365.
 28. Martina R, Jenkins D, Bujkiewicz S, Dequen P, Abrams K, GetReal Workpackage 1. The inclusion of real world evidence in clinical development planning. *Trials.* 29 août 2018;19(1):468.
 29. Lydersen S. Statistical review: frequently given comments. *Ann Rheum Dis.* févr 2015;74(2):323-5.
 30. Sormani MP, Signori A, Siri P, De Stefano N. Time to first relapse as an endpoint in multiple sclerosis clinical trials. *Mult Scler Houndmills Basingstoke Engl.* avr 2013;19(4):466-74.
 31. Friede T, Schmidli H. Blinded sample size reestimation with negative binomial counts in superiority and non-inferiority trials. *Methods Inf Med.* 2010;49(6):618-24.
 32. Friede T, Pohlmann H, Schmidli H. Blinded sample size reestimation in event-driven clinical trials: Methods and an application in multiple sclerosis. *Pharm Stat.* 2019;18(3):351-65.
 33. Sormani MP, Bruzzi P, Miller DH, Gasperini C, Barkhof F, Filippi M. Modelling MRI enhancing lesion counts in multiple sclerosis using a negative binomial model: implications for clinical trials. *J Neurol Sci.* 1 févr 1999;163(1):74-80.

Supplementary data

Searching terms used for reviewing

The following searching terms were entered in the searching tool of PubMed : (((((((((((("The New England journal of medicine"[Journal] OR "Lancet (London, England)"[Journal]) OR "JAMA"[Journal]) OR "The Lancet. Neurology"[Journal]) OR "JAMA neurology"[Journal]) OR "Brain : a journal of neurology"[Journal]) OR "Annals of neurology"[Journal]) OR "Neurology"[Journal]) OR "Journal of neurology, neurosurgery, and psychiatry"[Journal]) OR "Multiple sclerosis (Houndmills, Basingstoke, England)"[Journal]) OR "European journal of neurology"[Journal]) AND "multiple sclerosis"[All Fields]) AND ("clinical trials as topic"[MeSH Terms] OR ("clinical"[All Fields] AND "trials"[All Fields] AND "topic"[All Fields]) OR "clinical trials as topic"[All Fields] OR "trial"[All Fields]) AND ("2008/01/01"[PDAT] : "2019/12/31"[PDAT])).

Journal with IF \geq 4	Nb of screened articles	Nb of eligible articles	Nb of included articles
<i>Annals of Neurology</i>	45	3	3
<i>Brain</i>	14	0	0
<i>European Journal of Neurology</i>	49	0	0
<i>Journal of the American Medical Association (JAMA)</i>	5	0	0
<i>JAMA Neurology</i>	29	0	0
<i>Journal of Neurology, Neurosurgery, and Psychiatry</i>	59	0	0
<i>Lancet</i>	19	3	1
<i>Lancet Neurology</i>	70	6	6
<i>Multiple Sclerosis Journal</i>	440	2	1
<i>Neurology</i>	192	0	0
<i>New England Journal of Medicine</i>	20	9	8

Table S1. Identification of the articles included in the systematic review.

Exploratory endpoints.

Concerning the study characteristics, we picked up the study design (number of groups and allocation ratio) and its duration. We then collected the data necessary for the sample size estimation (SSE) : power, type I error (alpha risk), number of tails, expected annualized relapse rate (ARR) in the control group and in the intervention group, standard deviation of the ARR in both groups, expected dropout rate, published SSE, and the statistical test used for the SSE. When one of these informations was missing, it was recalculated using the available figures if possible (for example if the expected ARR of the intervention group was missing, it was calculated with the expected ARR in the control group and the difference expected). Data from the statistical analysis were also collected : statistical test used to compare the ARR, statistical significance (p value), observed ARR in control group and in intervention group.

Study	Power (%)	Alpha risk (%)	Tail(s)	ARR (control group)	ARR (treatment group)	Expected absolute difference	Dispersion (control group)	Dispersion (treatment group)	ARR reduction (%)	Dropout rate (%)
(19)	80	5	Two	1,1	0,77	0.33	0,6	0,6	30	NA
(14)	80	5	Two	0,8	0,56	0.24	NA	NA	30	15
(20)	95	5	NA	0,7	0,42	0.28	1,06	1,06	40	25
(11)	90	5	Two	0,55	0,33	0.22	0,9	0,9	40	13,4
(10)	90	5	Two	2,1	1,575	0.525	2,02	2,02	25	10
(12)	90	5	NA	0,49	0,343	0.147	NA	NA	30	NA
(21)	95	5	Two	0,74	0,555	0.185	0,626	0,626	25	NA
(15)	80	5	NA	0,33	0,2	0.13	NA	NA	39	21
(7)	90	5	Two	0,65	0,4875	0.1625	NA	NA	25	20
(22)	94	5	Two	0,61	0,456	0.154	NA	NA	25	23
(5)	90	5	Two	0,476	0,238	0.238	NA	NA	50	10
(9)	90	5	One	0,76	0,494	0.266	0,87	0,87	35	10
(13)	90	5	NA	0,35	0,245	0.105	NA	NA	30	15
(16)	93	5	Two	0,7	0,42	0.28	1,06	1,06	40	25
(17)	85	5	Two	0,43	0,2924	0.1376	NA	NA	32	10
(6)	94	5	Two	0,74	0,555	0.185	1,3*	1,3*	25	20
(8)	90	5	NA	0,27	0,2052	0.0648	NA	NA	24	21
(18)	80	10	Two	1,18	0,75	0.43	NA	NA	33	NA
(4)	84	5	Two	0,33	0,165	0.165	0,8	0,6	50	20

Table S2. Parameters needed for the SSE.

* Over-dispersion parameter for this study only. Dispersion parameter of other studies are standard deviation.

NA : Not available.

For Hauser et al. work (4), only the SSE of OPERA I is presented, as the statistical plan analyses of OPERA I and II were identical.

Study	Published planned sample size	Number of included patients	Method I Estimation with a normal distribution based on expected parameters and expected dispersion	Method II Estimation with a normal distribution based on expected parameters and observed dispersion	Method III Estimation with a negative binomial distribution based on expected parameters and observed dispersion	Method IV Estimation with a normal distribution based on observed parameters	Method V Estimation with a negative binomial distribution based on observed parameters
(19)*	180	181	159	-	-	-	-
(14)	300	130	-	274	628	177	66
(20)	1250	1272	1404	681	1947	1095	681
(11)	1275	1292	1206	1344	1929	2265	987
(10)	1290	1326	1032	78	3834	558	273
(12)*	980	1155	-	-	-	105078	-
(21)*	1080	1088	894	1380	2646	1629	1038
(15)	362	307	-	622	2170	5388	7144
(7)	1000	1106	-	-	-	-	-
(22)	1232	1430	-	2684	6628	2016	1416
(5)	600	621	-	804	855	726	669
(9)	1000	1008	608	-	-	-	-
(13)	1350	1404	-	6570	4230	2358	2979
(16)	1080	1083	1275	687	1614	1482	885
(17)	1500	1516	-	2904	3597	4647	3015
(6)	1110	1169	-	1689	5427	4746	7572
(8)	1800	1841	-	7212	2232	1056	700
(18)*	150	164	-	106	872	1340	1024
(4) ¹	800	821	992	810	692	1296	930
(4) ²	800	835	992	824	744	1326	994

Table S3. Comparison of the published planned sample size with the SSE following both normal and negative binomial distribution, based on expected and observed ARR parameters.

* These studies did not indicate an expected dropout rate in their statistical plan analysis. For these studies, our SS estimations were performed considering that there was no loss of follow-up. For all other studies, the SS includes the expected dropout rate announced by the authors.

For Hauser et al. (4) : ¹ corresponds to the OPERA I study, and ² to the OPERA II study. For both studies, we used the expected SD of the placebo group (0.8) to perform the SSE based on a normal distribution with expected parameters.

All SSE are calculated with two tails, except for Lublin et al. (9) (one tail).

Study	ARR control	CI 95%	Control group	ARR-1	CI 95%	Treatment group 1	ARR-2	CI 95%	Treatment group 2
(19)	1,05	NA	(IFN)	0,91	NA	(IFN+AZA)	0,73	NA	(IFN+AZA+cortico)
(14)	0,59	0.46 - 0.75	(IFN+placebo)	0,22	0.15-0.33	(IFN+méthylprednisolone)	-	-	-
(20)	0,4	0.34 - 0.47	(placebo)	0,18	0.15 to 0.22	(fingolimod 0,5)	0,16	0.13 to 0.19	(fingolimod 1,25)
(11)	0,33	0.26 - 0.42	(IFN)	0,16	0.12 to 0.21	(fingolimod 0,5)	0,2	0.16 to 0.26	(fingolimod 1,25)
(10)	0,33	0.29 - 0.38	(placebo)	0,14	0.12-0.17	(cladribine 3.5 mg/k)	0,15	0.12-0.17	(cladribine 5.25 mg/kg)
(12)	0,33	NA	(GA 20mg)	0,35	NA	(GA 40mg)	-	-	-
(21)	0,54	0.47 - 0.62	(placebo)	0,37	0.32-0.43	(teriflunomide 7mg)	0,37	0.31-0.44	(teriflunomide 14mg)
(15)	0,144	0.092 - 0.227	(IFN+placebo)	0,188	0.126 to 0.281	(IFN+simvastatine)	-	-	-
(7)	0,39	NA	(placebo)	0,3	NA	(laquinimod)	-	-	-
(22)	0,4	0.33 - 0.49	(placebo)	0,22	0.18-0.28	(DMF 2 daily)	0,2	0.16-0.25	(DMF 3 daily)
(5)	0,46	0.37 - 0.57	(placebo)	0,21	0.16-0.29	(daclizumab 150)	0,23	0.17-0.31	(daclizumab 300)
(9)	0,16	NA	(IFN)	0,11	NA	(GA 20mg)	0,12	NA	(IFN+GA 20mg)
(13)	0,505	0.418 - 0.609	(placebo)	0,331	0.280-0.392	(GA 40mg)	-	-	-
(16)	0,4	0.34 - 0.48	(placebo)	0,21	0.17 to 0.25	(fingolimod 0,5)	0,2	0.17 to 0.25	(fingolimod 1,25)
(17)	0,397	0.328 - 0.481	(placebo)	0,288	0.234-0.355	(PEG IFN 4W)	0,256	0.206-0.318	(PEG IFN 2W)
(6)	0,5	0.43 - 0.58	(placebo)	0,39	0.33 to 0.46	(teriflunomide 7mg)	0,32	0.27 to 0.38	(teriflunomide 14mg)
(8)	0,39	0.35 - 0.44	(IFN)	0,22	0.19-0.24	(daclizumab)	-	-	-
(18)	0,37	0.25 - 0.53	(GA 20mg)	0,25	0.17-0.37	(GA 20mg+estriol)	-	-	-
(4) ¹	0,29	0.24 - 0.36	(IFN)	0,16	0.12 to 0.20	(ocrelizumab)	NA	NA	NA
(4) ²	0,29	0.23 - 0.36	(IFN)	0,16	0.12 to 0.20	(ocrelizumab)	NA	NA	NA

Table S4. Observed ARR and dispersion in the studies, according to treatment group.
Dispersion is expressed with 95% CI.

CONCLUSION

CONCLUSION

Dans ce travail de thèse, nous avons conduit une revue systématique de la littérature sur le calcul du Nombre de Sujets Nécessaires (NSN) à inclure dans les essais cliniques randomisés contrôlés de phase III et IV étudiant les traitements de fond de la sclérose en plaques de forme récurrente-rémittente. Les résultats de ce travail montrent que ce calcul n'est pas optimal. En effet, le Taux Annualisé de Poussées (TAP), fréquemment utilisé comme critère de jugement principal, suit une distribution obéissant à la loi des événements rares. La majorité des essais cliniques étudiant le TAP ne prend pas en considération ce type de distribution lors du calcul du NSN, alors qu'ils en tiennent compte lors de l'analyse statistique.

Dans un deuxième temps, nous avons modélisé le calcul du NSN à partir des données extraites de chaque étude avec différentes méthodes. Les simulations de calcul du NSN ont donné des résultats très variables, en fonction du TAP et de la dispersion utilisés, mais montre en moyenne un NSN moins important avec une distribution binomiale négative, qu'avec un modèle gaussien. Ces résultats sont très probablement expliqués par la complexité de la formule de calcul basée sur une distribution BN. Comme il est difficile de prévoir ces variations, l'utilisation d'un modèle BN semble obligatoire pour optimiser le nombre de sujet nécessaire.

Clermont-Ferrand, le 11/3/20

Pierre CLAVELOU
Doyen - Directeur



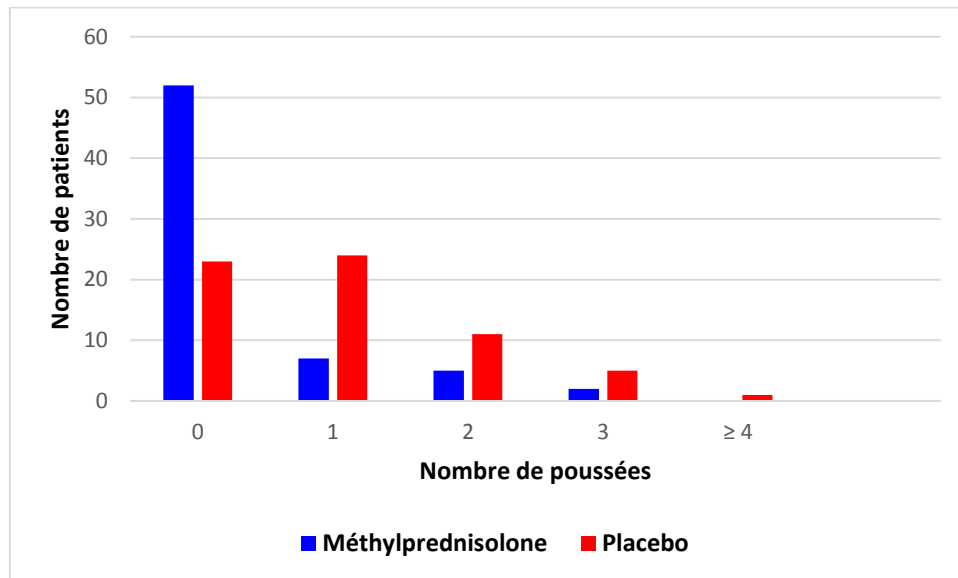
Clermont-Ferrand, le 11/3/20

Le Président du Jury

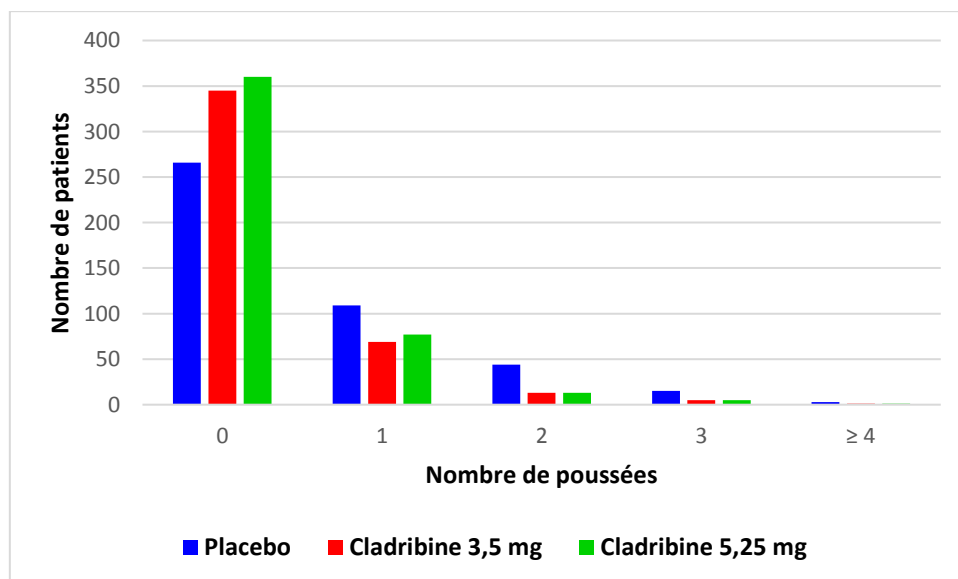
C.H.U. Gabriel Montpied
Clermont-Ferrand
Professeur P. CLAVELOU
Service de neurologie
N° RPPS 10003884896
N° FINES 630780989

ANNEXES

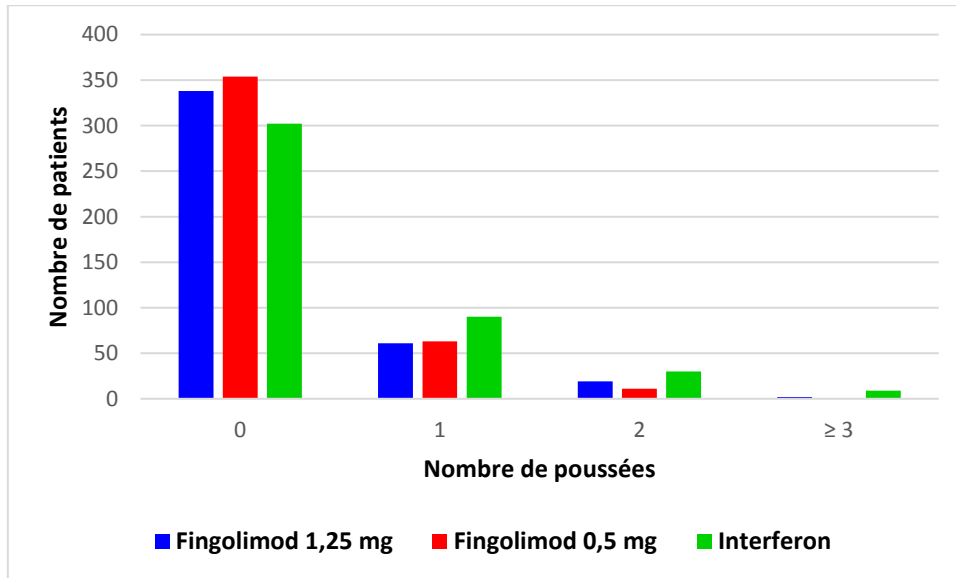
Annexe I : Répartition du nombre de poussées en fonction du groupe de traitement dans les études le précisant.



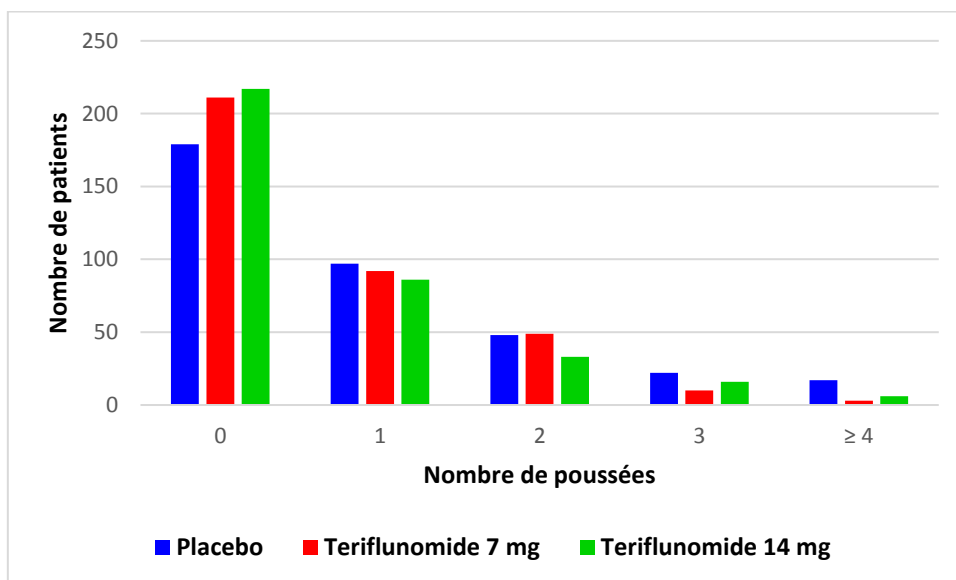
P. S. Sorensen et al., NORdic trial of oral Methylprednisolone as add-on therapy to Interferon beta-1a for treatment of relapsing-remitting Multiple Sclerosis (NORMIMS study): a randomised, placebo-controlled trial. *Lancet Neurol.* 8, 519–529 (2009).



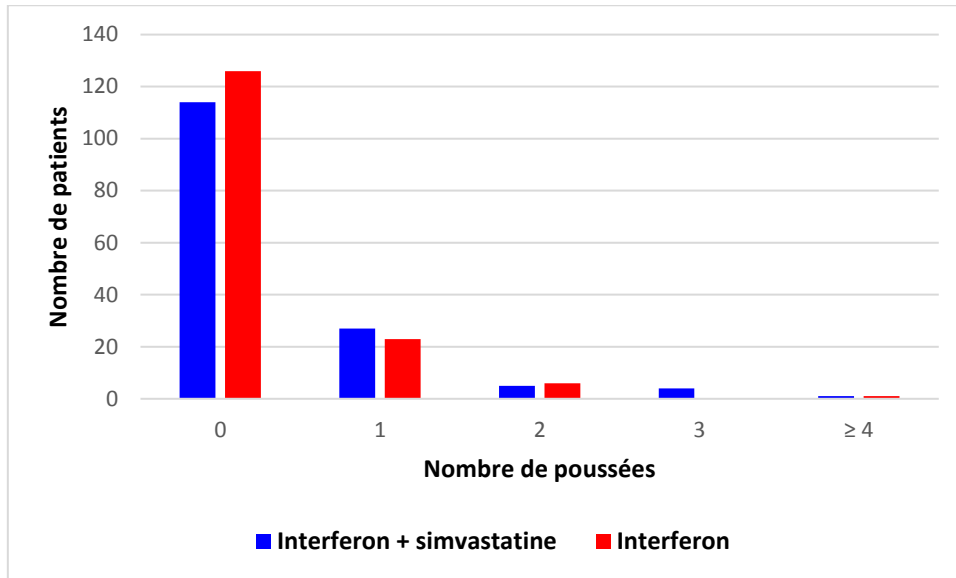
G. Giovannoni et al., A placebo-controlled trial of oral cladribine for relapsing multiple sclerosis. *N. Engl. J. Med.* 362, 416–426 (2010).



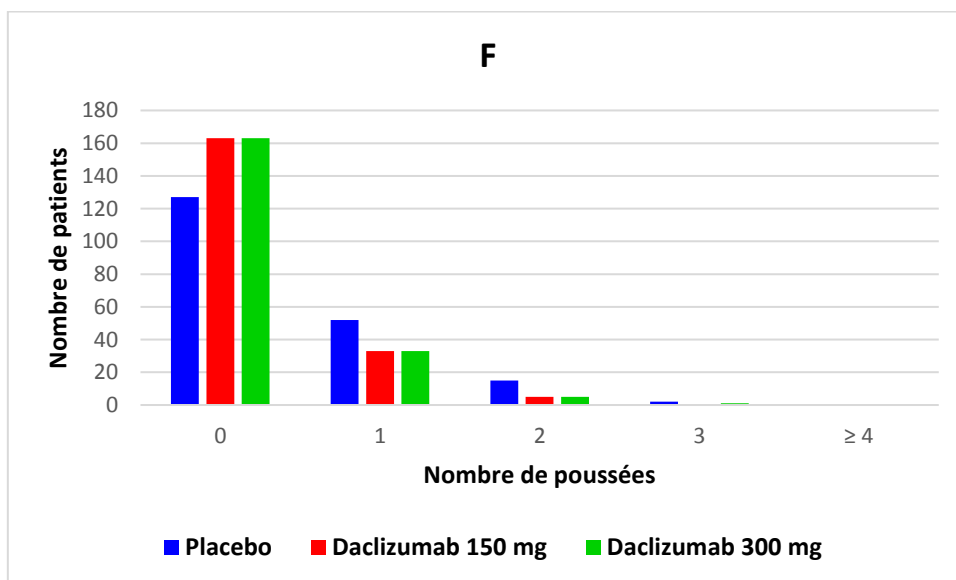
J. A. Cohen *et al.*, Oral fingolimod or intramuscular interferon for relapsing multiple sclerosis. *N. Engl. J. Med.* **362**, 402–415 (2010).



P. O'Connor *et al.*, Randomized trial of oral teriflunomide for relapsing multiple sclerosis. *N. Engl. J. Med.* **365**, 1293–1303 (2011).



P. S. Sorensen et al., Simvastatin as add-on therapy to interferon β -1a for relapsing-remitting multiple sclerosis (SIMCOMBIN study): a placebo-controlled randomised phase 4 trial. *Lancet Neurol.* 10, 691–701 (2011).



R. Gold et al., Daclizumab high-yield process in relapsing-remitting multiple sclerosis (SELECT): a randomised, double-blind, placebo-controlled trial. *Lancet Lond. Engl.* 381, 2167–2175 (2013).

Annexe II : Formule utilisée pour le calcul du NSN selon une loi binomiale négative. Source : PASS Sample Size Software (dérivé de Zhu and Lakkis, 2014).

Calcul du NSN :

$$n \geq \frac{(z_{\alpha/2}\sqrt{V_0} + z_{\beta}\sqrt{V_A})^2}{(\log(RR))^2}$$

$$\text{Avec } V_0 = \frac{1}{\mu_t} \left(\frac{1}{\tilde{\lambda}_1} + \frac{1}{R\tilde{\lambda}_2} \right) + \frac{(1+R)\kappa}{R}$$

$$\text{Et } V_A = \frac{1}{\mu_t} \left(\frac{1}{\lambda_1} + \frac{1}{R\lambda_2} \right) + \frac{(1+R)\kappa}{R}$$

κ : paramètre de dispersion binomial négatif

n : NSN par groupe

R : ratio d'allocation (n_2 / n_1)

RR : risk ratio (λ_2 / λ_1)

λ_1 et λ_2 : taux annualisé de poussée dans chaque groupe

μ_t : durée totale d'exposition pour l'ensemble des sujets

Estimation de κ à partir de l'intervalle de confiance à 95% d'une étude précédente :

Admittedly, the hardest value to determine among those required for these sample size and power calculations is the value for κ , the negative binomial dispersion parameter. If a suitable value for κ is not known, then you can estimate κ from a similar study that was analyzed using Poisson regression.

Given a Poisson mean event rate λ and an overdispersion factor ϕ , estimated from a similar Poisson regression study, the relationship between λ , ϕ , and κ is

$$\phi = 1 + \kappa\lambda$$

such that an estimate of κ can be calculated using overall estimates of λ and ϕ as

$$\hat{\kappa} = \frac{\hat{\phi} - 1}{\hat{\lambda}}$$

If $\hat{\phi}$ (the estimate for the overdispersion factor) is not directly reported in the previous study, Zhu and Lakkis (2014) suggest on page 384 that the Poisson overdispersion factor can be estimated from the total number of events, Y , the total exposure time, T , and the standard error, $SE(\log \hat{\lambda})$, as

$$\hat{\phi} = \frac{\text{Var}(Y)}{\hat{\lambda}T} \approx T\hat{\lambda} \left(SE(\log \hat{\lambda}) \right)^2$$

since for overdispersed Poisson

$$\text{Var}(Y) = \phi\lambda T$$

The maximum likelihood estimator of the event rate of the Poisson distribution, λ , is

$$\hat{\lambda} = \frac{Y}{T}$$

The standard error, $SE(\log \hat{\lambda})$, can be determined by back-calculating from reported $100(1 - \alpha)\%$ confidence interval endpoints for λ from Poisson regression as

$$SE(\log \hat{\lambda}) = \frac{\left(\frac{\log(\text{Upper Bound}) - \log(\text{Lower Bound})}{2} \right)}{z_{1-\alpha/2}}$$

Finally, from a previous two-group Poisson regression study, κ can be calculated from the estimated Poisson event rates $\hat{\lambda}_1$ and $\hat{\lambda}_2$ and the estimated Poisson overdispersion factors $\hat{\phi}_1$ and $\hat{\phi}_2$ from each group as

$$\hat{\kappa} = \frac{\left(\frac{\hat{\phi}_1 + \hat{\phi}_2}{2} \right) - 1}{\left(\frac{\hat{\lambda}_1 + \hat{\lambda}_2}{2} \right)}$$

SERMENT D'HIPPOCRATE

SERMENT D'HIPPOCRATE

(Conseil national de l'ordre des médecins)

Au moment d'être admis(e) à exercer la médecine, je promets et je jure d'être fidèle aux lois de l'honneur et de la probité.

Mon premier souci sera de rétablir, de préserver ou de promouvoir la santé dans tous ses éléments, physiques et mentaux, individuels et sociaux.

Je respecterai toutes les personnes, leur autonomie et leur volonté, sans aucune discrimination selon leur état ou leurs convictions. J'interviendrai pour les protéger si elles sont affaiblies, vulnérables ou menacées dans leur intégrité ou leur dignité. Même sous la contrainte, je ne ferai pas usage de mes connaissances contre les lois de l'humanité.

J'informerai les patients des décisions envisagées, de leurs raisons et de leurs conséquences. Je ne tromperai jamais leur confiance et n'exploiterai pas le pouvoir hérité des circonstances pour forcer les consciences.

Je donnerai mes soins à l'indigent et à quiconque me les demandera. Je ne me laisserai pas influencer par la soif du gain ou la recherche de la gloire.

Admis(e) dans l'intimité des personnes, je tairai les secrets qui me seront confiés. Reçu(e) à l'intérieur des maisons, je respecterai les secrets des foyers et ma conduite ne servira pas à corrompre les mœurs. Je ferai tout pour soulager les souffrances. Je ne prolongerai pas abusivement les agonies. Je ne provoquerai jamais la mort délibérément.

Je préserverai l'indépendance nécessaire à l'accomplissement de ma mission. Je n'entreprendrai rien qui dépasse mes compétences. Je les entretiendrai et les perfectionnerai pour assurer au mieux les services qui me seront demandés.

J'apporterai mon aide à mes confrères ainsi qu'à leurs familles dans l'adversité.

Que les hommes et mes confrères m'accordent leur estime si je suis fidèle à mes promesses ; que je sois déshonoré(e) et méprisé(e) si j'y manque.

SERMENT D'HIPPOCRATE

(Faculté de Médecine de Clermont-Ferrand)

En présence des Maîtres de cette FACULTE et de mes chers CONDISCIPLES, je promets et je jure d'être fidèle aux lois de l'Honneur et de la Probité dans l'exercice de la Médecine.

Je donnerai mes soins gratuits à l'indigent et je n'exigerai jamais un salaire au-dessus de mon travail. Admis dans l'intérieur des maisons, mes yeux ne verront pas ce qui s'y passe, ma langue taira les secrets qui me seront confiés et mon état ne servira pas à corrompre les mœurs ni à favoriser le crime.

Respectueux et reconnaissant envers mes MAÎTRES, je rendrai à leurs enfants l'instruction que j'ai reçue de leurs pères.

Que les HOMMES m'accordent leur estime si je suis fidèle à mes promesses. Que je sois couvert d'OPPROBRE et méprisé de mes confrères si j'y manque.

TITRE : Calcul du nombre de sujets nécessaires dans les essais cliniques randomisés contrôlés de phase III et IV étudiant les traitements de fond de la sclérose en plaques de forme récurrente-rémittente : une revue systématique

RESUME :

Contexte : Dans les essais cliniques randomisés contrôlés, le calcul du Nombre de Sujets Nécessaires (NSN) à inclure dans l'étude est une étape importante de la planification de l'analyse statistique. Le calcul doit être adapté à la distribution de la variable étudiée. Dans la Sclérose En Plaques (SEP), le Taux Annualisé de Poussées (TAP) suit une distribution binomiale négative (BN).

Objectifs : Déterminer si le calcul du NSN et l'analyse du TAP dans les essais cliniques randomisés contrôlés de phase III et IV étudiant les traitements de fond de la SEP de forme récurrente-rémittente (RR) sont en accord avec la distribution du TAP, et comparer le calcul du NSN selon plusieurs distributions statistiques.

Méthodes : Une revue systématique de la littérature a été conduite, incluant les essais cliniques randomisés contrôlés portant sur la SEP-RR et ayant comme critère de jugement principal le TAP, publiés depuis 2008 dans des journaux scientifiques ayant un Impact Factor supérieur ou égal à 4. L'adéquation du calcul du NSN à la distribution du TAP a été vérifiée. Nous avons également comparé le NSN basé sur une distribution normale et BN, en utilisant les données recueillies dans le cadre des études.

Résultats : Dix-neuf articles correspondant à 20 études ont été retenus. Neuf d'entre eux indiquaient la méthode utilisée pour le calcul du NSN. Parmi ceux-ci, seuls 4 utilisaient une formule basée sur une distribution binomiale négative ou dérivée d'un modèle de Poisson. L'analyse du TAP était en revanche basée sur ces 2 modèles pour la quasi-totalité des études. A partir de paramètres identiques, la modélisation du calcul du NSN avec les différentes méthodes a donné des résultats très variables, en fonction du TAP et de la dispersion utilisés. Néanmoins, le NSN était en moyenne moins important lors de l'utilisation d'une distribution BN qu'à partir d'un modèle gaussien.

Conclusions : Notre étude a montré que le calcul du NSN dans les essais randomisés contrôlés de la SEP étudiant le TAP est rarement réalisé avec la bonne méthode. La complexité de la formule du modèle binomial négatif engendre une variabilité du NSN, rendant son approximation impossible par un modèle gaussien. Il semble donc obligatoire d'utiliser un calcul du NSN basé sur un modèle BN pour effectuer des analyses statistiques rigoureuses méthodologiquement.

MOT-CLES : Revue systématique, Sclérose en plaques récurrente-rémittente, Nombre de sujets nécessaire, Essai randomisé contrôlé, Binomial négatif