Multiple Sclerosis and Related Disorders 35 (2019) 142–149

Contents lists available at ScienceDirect

Multiple Sclerosis and Related Disorders

journal homepage: www.elsevier.com/locate/msard

Review article

Observational designs in clinical multiple sclerosis research: Particulars, practices and potentialities

Peter Joseph Jongen[a,b,*]

[a] University Medical Centre Groningen, Department of Community and Occupational Medicine, Antonius Deusinglaan 1, 9713 AV, Groningen, the Netherlands
[b] MS4 Research Institute, Nijmegen, the Netherlands

ARTICLE INFO

Keywords:
Multiple sclerosis
Observational
Effectiveness
Disease-modifying drug
Pharmacotherapy

ABSTRACT

Observational studies investigate a wide range of topics in multiple sclerosis research. This paper presents an overview of the various observational designs and their applications in clinical studies. Observational studies are well suited for making discoveries and assessing new explanations of phenomena, but less so for establishing causal relationships, due to confounding by indication (selection bias), co-morbidity, socio-economic or other factors. Whether observational findings are demonstrative, indicative or only suggestive, depends on the research question, whether and how the design fits this question, analytical techniques, and the quality of data. Observational studies may be cross-sectional vs. longitudinal, and prospective vs. retrospective. The term 'retrograde' is proposed to explicate that cross-sectional studies may obtain data that cover (long) preceding periods.

Case reports and case series are usually based on accidental observations or routinely collected data. Cross-sectional studies, by simultaneously assessing clinical phenomena and external factors, enable the discovery and quantification of associations. In ecological studies the unit of analysis is population or group, and relationships on patient level cannot be established. A cohort study is a longitudinal study that investigates patients with a defining characteristic, e.g. diagnosis or specific treatment, by analyzing data acquired at various intervals. Prospective cohort studies use (some) data that are not yet available at the time the research is conceived, whereas in retrospective studies the data already exist. In a case-control study a representative group of patients with a specific clinical feature is compared with controls, and the frequencies at which an external factor, e.g. infection, has occurred in each group is compared; in a nested case-control study controls are drawn from a fully known cohort. Randomized controlled trial (RCT)-extension studies are informative because, due to RCT randomization, they are free from confounding by indication. Patient or disease registries are organised systems for the long-term collection of uniform data on a population that is defined by a particular disease, condition or exposure, with the purpose to study changes over time.

In pharmacotherapeutic research, accidental observations of unexpected beneficial effects may lead to further research into a drug's efficacy in other conditions. Uncontrolled phase 1 studies investigate safety and dosing aspects. Observational studies are alternatives to RCTs when these are not feasible for ethical or practical reasons. Phase 4 observational studies play a crucial role in the evaluation of the effectiveness of treatments in daily practice, the validation of RCT-based side effect profiles, and the discovery of late occurring or rare, potentially life-threatening side effects. Combinations of multidisciplinary longitudinal data bases into large data sets enable the development of algorithms for personalized treatments.

To improve the reporting of observational findings on treatment effectiveness, it is proposed that abstracts define the research question(s) the study was meant to answer, study design and analytical methods, and identify and quantify the patient population, treatment of interest, relevant outcomes and the study's strengths and limitations. The development of guidelines for Strengthening the Reporting of Observational Studies in Effectiveness Research (STROBER), as an extension of the guidelines used in epidemiology, is wanted.

* Corresponding author at: University Medical Centre Groningen, Department of Community and Occupational Medicine, Antonius Deusinglaan 1, 9713 AV, Groningen, the Netherlands.
E-mail addresses: p.j.h.jongen@rug.nl, ms4ri@kpnmail.nl.

https://doi.org/10.1016/j.msard.2019.07.006
Received 27 March 2019; Received in revised form 3 July 2019; Accepted 19 July 2019

1. Introduction

The importance of observational studies in multiple sclerosis (MS) research has long been under-appreciated by neurologists and clinical researchers (Feigin and Howard, 2008). In the last decades, due to an increasing number of randomised controlled trials (RCTs) on the efficacy of disease-modifying drugs (DMDs), the RCT design became to be perceived as virtually the only valid method in therapeutic research (Feigin and Howard, 2008). In fact, however, a considerable number of theoretical and practical questions can be answered only with observational data (Feigin and Howard, 2008).

In recent years the appraisal of observational studies in clinical MS research has increased (Trojano et al., 2007), especially since neurologists have come to realize that discussions regarding the effectiveness of DMD treatments can be resolved only by using high-quality data acquired via observations in daily practice (Kalincik and Butzkueven, 2016; Sormani and Bruzzi, 2015; Trojano et al., 2017). Thus, a PubMed search using the entries 'multiple sclerosis' and 'observational' showed that in the years 2008–2018 the number of articles had increased from 21 to 160 per year. Accordingly, a common understanding of the evidence from observational studies is becoming a basis for supporting clinical decision-making.

However, the critical valuation and interpretation of observational reports by clinical neurologists may be hampered by specific methodological aspects of study designs they are not sufficiently familiar with (Hernan et al., 2004; Kalincik and Butzkueven, 2016; Trojano et al., 2017).

In view of the above, this article aims to present an overview of observational study designs used in clinical MS research, and of those pharmacotherapeutic research questions that can be answered only or mainly by observational data. The latter regard the effectiveness of treatments, validation of RCT-based side effect and safety profiles, identification of late or rare side effects, and alternatives to RCTs where these are not feasible.

2. Epistemological aspects

2.1. Observations in medical science

An observation may be defined as 'an act of recognizing and noting a fact or occurrence often involving measurement with instruments', as well as 'a judgment on or inference from what one has observed' (Merriam-Webster, 2018). Accordingly, observational studies are studies that draw inferences in settings where the independent variables are not under the control of researchers (non-interventional studies) or where there is no randomization (non-randomized interventional studies) (Vandenbroucke, 2008).

The scientific status of observational studies largely depends on the research question. Thus, regarding the establishment of intended treatment effects, the RCT has by far the strongest evidential value (Feigin and Howard, 2008; Vandenbroucke, 2008), whereas observational studies have serious interpretational difficulties in this respect. An association not necessarily reflects causality, nor the direction of causality, as it may be coincidental or mediated by a third factor (Vandenbroucke, 2008). Very rarely do observational reports provide sufficient evidence for a treatment effect (Vandenbroucke, 2008; Smith and Pell, 2003), in contrast to them being well suited for making discoveries and assessing new explanations of phenomena (Vandenbroucke, 2008; Frieden, 2017).

2.2. Confounding

When assessing the effect of a treatment on an outcome via observations the possibility of confounding by other factors must be considered (Kyriacou and Lewis, 2016; Rothman, 2008). To be qualified as confounding a variable has to meet three criteria (Rothman, 2008): it must be an independent causal (or a surrogate for a casual) factor for the outcome, must be associated with the exposure, and cannot be an intermediate variable between treatment and outcome (Kyriacou and Lewis, 2016; Mansournia et al., 2017; Hernan et al., 2004).

Confounding by indication (selection bias) occurs when the indication for choosing a specific treatment also affects the outcome; e.g. patients with a high annualized relapse rate (ARR) are likely to be treated with high-efficacy DMDs, and when comparing treatment effects in non-randomized studies, the treatment with high-efficacy DMDs may appear to result in outcomes equal to those of low-efficacy DMDs in patients with low disease activity (Kyriacou and Lewis, 2016). Another example of selection bias is the preferential inclusion in RCTs of patients with high disease activity when there is no standard treatment, and the inclusion of patients with low disease activity when standard treatment is available (Vandenbroucke, 2008).

Other examples of confounders are clinical conditions that exist before or during a patient's treatment, as they may affect the outcome and influence the type of treatment (Kyriacou and Lewis, 2016); and socio-economic factors, as they may affect the type of treatment and the availability of overall medical care (Kyriacou and Lewis, 2016).

2.3. Technical terms

Observational designs can be characterized by technical terms. Four are discussed here to prevent misunderstandings and promote an unambiguous use: cross-sectional, longitudinal, prospective and retrospective. A fifth term is introduced.

'Cross-sectional' are studies that focus on describing a state or phenomenon at a fixed or indefinite time (Bailar III JC, 1986). Thus, a study of the pharmacokinetics of a DMD describes characteristics in an indefinite time; and a study on the association between MS diagnosis and genetic make-up investigates phenomena that are fixed in time (diagnosis) and indefinite (genetic make-up). Notably, cross-sectional studies may obtain data that cover preceding periods, of months or even years (Bailar III JC, 1986). Thus, health-related quality of life (HRQoL) questionnaires typically ask about the previous one to four weeks; and to calculate the Medication Possession Ratio, pharmacy data are usually obtained over a period of months or years (Bailar III JC, 1986). To explicate this potential backward aspect of data collection in cross-sectional studies, we propose the term 'retrograde' data collection.

'Longitudinal' studies analyse processes over time to investigate changes (Bailar III JC, 1986), and therefore data acquisitions are performed at various intervals through time. The qualification 'longitudinal' does not depend directly on the length of the time period covered by the observations (Bailar III JC, 1986). E.g. a study on the cardiovascular effects of a sphingosine 1-phosphate-receptor-1 modulator would be considered longitudinal even though the observations in patients may last an hour or less (Bailar III JC, 1986).

Prospective longitudinal studies collect or evaluate at least some data that are not yet available at the time the research is conceived, whereas retrospective longitudinal studies collect or evaluate data that already exist at the time that the study is conceived.

In fact, two or more observational designs may be combined in one study. E.g. in a registry-based study on the association between DMD treatment and HRQoL, existing data may be analysed (retrospective longitudinal aspect) in combination with cross-sectional data (retrograde cross-sectional aspect), and possibly a prospective follow-up.

The quality of observational data relate in part to the methodological aspects the above terms refer to. E.g. retrospective studies may analyse data that have been collected in daily practice without research intention, whereas prospective studies are usually performed on the basis of well-defined questions and protocols, and possibly with the application of quality assurance measures. Moreover, in MS research, due to patients´ cognitive impairment, retrograde acquisition of patient-reported outcomes (PROs) may negatively affect data quality.

3. Observational study designs

3.1. Case reports and case series

These are usually based on accidental observations in clinical practice or routinely collected clinical data (Thrift, 2009). Although the scientific community will only rarely consider data from case reports or case series to be a sufficient explanation of intended changes, there are historical examples of treatments with such dramatic effects – e.g. insulin for diabetes – that they were not further investigated in RCTs (Glasziou et al., 2007; Vandenbroucke, 2008).

The evidential strength of case reports and series is high when pertaining to unintended phenomena. Two case reports of progressive multifocal leukoencephalopathy (PML) in relapsing remitting (RR) MS patients treated with natalizumab in the extension phase of a RCT – in combination with a PML report in a natalizumab-treated patient with Crohn's disease – convinced neurologists, researchers and health authorities of the causal link between natalizumab treatment and PML (Kleinschmidt-DeMasters and Tyler, 2005; Langer-Gould et al., 2005; Van Assche et al., 2005). The acceptance of causality was facilitated by the background knowledge: PML had been known to occur almost exclusively in patients with severe immune deficiency.

3.2. Cross-sectional studies

In cross-sectional studies the simultaneous assessment of a clinical phenomenon and an external factor enables the discovery and quantification of associations (Thrift, 2009). Practical advantage of the cross-sectional design is that studies can be performed relatively easy (Thrift, 2009). However, to evaluate the causality of associations special statistical techniques, like propensity score matching, are required. E.g. a cross-sectional study in two groups of 50 RRMS patients, each treated for at least one year with intramuscular interferon-beta (INF-β)-1a or its biosimilar product, found no statistically significant differences regarding HRQoL (Hatam et al., 2016). Although the groups did not differ in demographic or disease characteristics, the results do not justify the conclusion that the effect of the two DMDs on HRQoL does not differ, as statistical procedures correcting for the absence of randomization were not applied (Hatam et al., 2016). The interpretation of an association may be further complicated when it is unknown whether or not the exposure to a potential risk factor has preceded the occurrence of the phenomenon (Thrift, 2009).

3.3. Ecological studies

In ecological studies the unit of analysis is population or group, and relationships on patient level, e.g. between exposure and effect, cannot be established (Thrift, 2009). Populations or groups are compared between countries, or within one country at different time points (Thrift, 2009). Generally, data are used that have been obtained for other purposes, which explains why the interpretation of the findings may pose difficulties (Thrift, 2009).

3.4. Cohort studies

A cohort study is a longitudinal observational study that investigates a group of patients who share a defining characteristic, e.g. diagnosis or specific treatment, by analyzing data that have been acquired at various intervals through time.

3.4.1. Prospective cohort studies

Prospective cohort studies collect or evaluate at least some data that are not yet available at the time the research is conceived. Typically, patients without the characteristic under investigation are followed in time and are assessed for the occurrence of the characteristic. Examples in MS are: conversion from clinically isolated syndrome to RRMS, from RRMS to secondary progressive MS, or the occurrence of a relapse, disability progression or a specific adverse event. In general, after evaluation of patients who have been exposed and who have not been exposed to an external factor – e.g. treatment –, the comparison of the frequencies in both groups enables the assessment of a possible relationship between the factor and the characteristic (Thrift, 2009). The advantages of prospective cohort studies are that all cases can be ascertained, that data collection does not depend on the patients' memory or hospitals' participation, that all patients are potentially assessed in the same way, and that multiple outcomes can be measured (Thrift, 2009). Thus, MS patients with and without DMD treatment can be followed up and associations can be established between disability progression and treatment, provided that one corrects for the absence of randomization (Trojano et al. 2007, 2009) .

Observational studies of intended changes in patients treated with a specific (class of) DMD(s) are a special type of prospective cohort studies. Although no firm conclusion can be drawn on the causal relationship between treatment and outcomes, these studies may quantify on a group level changes occurring during treatment, and thus inform about the chance of such changes in other groups in the same or similar populations (Jongen et al., 2010a,b, 2014).

3.4.2. Retrospective cohort studies

Retrospective cohort studies collect or evaluate data that are existing at the time the research is conceived. Thus, to study the effect of DMD treatment on progression to EDSS 3.0 and 6.0, and the relationship between treatment effect and baseline disability, a hospital data base of 3060 MS patients with annual clinical assessments was analyzed, applying Cox regression analysis adjusted for propensity score and immortal time bias (Cocco et al., 2015). In comparison with untreated patients, the risk of EDSS 3.0 was 94% lower in patients receiving immunomodulating therapy, and the risk of EDSS 6.0 was 86% lower; moreover, in comparison with untreated patients, the risk of EDSS 6.0 was 91% lower in patients who received immunomodulation before EDSS 3.0, and 75% lower in patients who received such treatment after EDSS 3.0 (Cocco et al., 2015). However, as this study was performed in a single academic center the generalizability of the results may be limited (Cocco et al., 2015).

In general, retrospective cohort studies are less informative than prospective ones, as data may be incomplete (Kappos et al., 2015b). Especially in long-term studies, patient assessments may be relatively infrequent, follow-up intervals may vary, and diverse methods may have been used for assessment (Ford et al., 2010). Other aspects of retrospective cohort designs may also determine whether conclusions are informative. E.g., to assess the long-term effectiveness of GA treatment in clinical practice conditions, a retrospective cohort study included 149 RRMS patients treated with GA for at least five years, and defined long-term effectiveness as absence of disability progression for at least five consecutive years (Arnal-Garcia et al., 2014). About 75% of the patients had no disability progression for at least five consecutive years. However, as the study selectively included patients who were 5-year treatment-persistent, and did not inform about non-persistence in the first five years, the practical relevance of the findings are limited (Arnal-Garcia et al., 2014).

3.4.3. Registry-based cohort studies

Patient or disease registries are organised systems for the long-term collection of uniform data on a population that is defined by a particular disease, condition or exposure, with the purpose to study changes over time. Accordingly, MS registries may be general, aiming at all MS patients in a defined area, or may focus on specific subgroups, e.g. pregnant patients (Alwan et al., 2015), patients exposed to DMD treatment (Ghezzi et al., 2015) or those with certain side effects (Prosperini et al., 2016). Registries are typically used for large-scale prospective and retrospective cohort studies. Most MS registries are nationwide population-based quality registers, enabling health

authorities to monitor the quality and effects of DMD treatments (Flachenecker et al., 2014). Examples of registry types and their scientific potential are presented here.

In addition to the epidemiological Danish MS Registry, there is the obligatory Danish MS Treatment Register for patients who have received DMD treatments (Koch-Henriksen et al., 2015; Magyari et al., 2016; Boesen et al., 2018). Notably, analysis of the latter registry's data of 496 patients treated with fingolimod showed that 3.9% required prolonged cardiac monitoring due to bradycardia or second-degree atrioventricular block type I, a rate similar to that reported in RCTs (Voldsgaard et al., 2017).

The voluntary Swedish national MS registry exerts nationwide pharmacological surveillance (Andersen, 2012; Hillert and Stawiarz, 2015; Granqvist et al., 2018). Interestingly, a recent study of 2477 patients from this registry, followed up to 15 years, revealed that starting DMD treatment within six months after the onset of disease is associated with a 36% lower risk (HR 0.74, $p = 0.010$) of full-time disability pension in comparison with starting treatment after 18 months (Landfeldt et al., 2018).

The North American Research Committee on MS registry is a voluntary self-reported registry, collecting clinical and socio-demographic information through semi-annual surveys (Salter et al., 2014). Recent studies showed that in MS patients satisfaction with treatment of bladder, bowel and sexual problems is low (9341 patients) (Wang et al., 2018; Fitzgerald et al., 2018), and that stigma is related to depression (5369 patients) (Cadden et al., 2018).

The MSBase registry is an international neurologists-driven project to study MS and other neuro-immunological disorders. The registry contains over 52,000 patient records from 34 countries, making it the largest repository of longitudinal, observational MS patient data worldwide (Butzkueven et al., 2006). A recent analysis showed that among RRMS patients initial treatment with fingolimod, alemtuzumab or natalizumab was associated with a lower risk of conversion to secondary progressive MS vs. initial treatment with GA or INF-β (Brown et al., 2019).

3.5. (Nested) case-control studies

3.5.1. Case-control studies

In a case-control study a representative group of patients with a specific clinical feature is compared with an otherwise similar group of controls (Thrift, 2009). The frequencies at which an external factor (e.g. treatment, risk factor exposure) has occurred in each group is compared, and the result is expressed as an odds ratio (OR) (Thrift, 2009). The case-control design is particularly advantageous when studying rare outcomes because of the savings in cost and time that would otherwise be required to study all the controls in the population from which the cases are selected (Essebag et al., 2003). The underlying assumption of a case-control study is that cases and controls are random samples selected from the same source population. However, when controls are selected from other clinics or neighborhoods of cases, selection bias may be introduced (Essebag et al., 2003). A recent case-control study on the prevalence of malignancies in MS used a postal survey of a large cohort of MS patients and data from consecutive outpatients as controls (Moisset et al., 2017). After propensity score matching for age, gender, smoking history and alcohol consumption, it was found that 7.32% of MS patients had ever presented with a cancer vs. 12.63% of controls (Moisset et al., 2017).

3.5.2. Nested case-control studies

A variant of the case-control design is the nested case-control study in which controls are drawn from the population in a fully enumerated cohort for which information on all members can be obtained (Essebag et al., 2003). This makes it easier to satisfy the assumption that cases and controls are random samples of the same study population (Essebag et al., 2003). A common definition is to also demand 'density' (or 'risk set') sampling of controls, i.e. that controls are selected at the time a person becomes a case, among subjects currently at risk and not yet having developed the outcome. This clever sampling means that one estimates the odds of being an incident case, the OR being a good approximation of the incidence rate. Nested case-control studies are common today, in particular nested in registries, as they overcome several of the potential weaknesses of traditional case-control studies. Thus, the nested case-control design is frequently used when the exposure of interest is difficult or expensive to obtain and when the outcome is rare. Recently, a nested case-control study assessed the odds of previous INF-β exposure for potential adverse events in RRMS patients, and found that INF-β was associated with a 1.83 (95% CI 1.16–2.89) and 1.55 (95%CI 1.18–2.04)-fold increase in the risk of stroke and migraine, respectively, and a 1.33 (95%CI 1.13–1.56) and 1.32 (95% CI 1.01–1.72)-fold increase in depression and hematologic abnormalities (de Jong et al., 2017; Jongen, 2017b).

Case-control studies can also use historical data to compare with prospective observations. Recently, the medium-term effectiveness and cost-effectiveness of INF-β and GA treatment were studied in the UK by using the UK Multiple Sclerosis Risk Sharing Scheme cohort and a natural history comparator (Palace et al., 2015). Combining observational data on the progression ratio (treated vs. untreated, measured in EDSS score and utility) with modelling techniques provided evidence that in RRMS patients the effects of INF-β and GA treatment on disability are maintained and cost-effective over 6 years (Palace et al., 2015).

3.6. Randomized controlled trial follow-up studies

RCT follow-up studies, in which patients in the experimental treatment arm are compared with those having received placebo or standard treatment, are a special type of observational study. These studies are informative because, even though they compare only early vs. delayed treatment, due to the randomization in the RCT phase they are free from confounding by indication (Sormani and Bruzzi, 2015). RCT follow-up studies may be prospective, designed as RCT-extension studies; retrospective, when the study is limited to data that have been documented in health records in daily practice; or cross-sectional, when a single assessment is performed at a fixed time point possibly with retrograde data collection (Ford et al., 2010, 2006; Kappos et al., 2015b). In general, cross-sectional follow-up studies investigating long-term effectiveness of treatments may be hampered by patients lost to follow-up or not willing to participate, or centres not willing to cooperate. In all, results from RCT follow-up studies indicate that RRMS patients can experience sustained benefit over many years from early DMD therapy, in comparison with patients whose treatment is delayed (Kappos et al., 2015b, 2006).

3.7. Combinations of observational data

There is a trend to combine data from various data bases into large data sets, as it is thought that information from multidisciplinary and longitudinal assessments – including clinical, functional and PRO measures – will facilitate the development of personalized treatment and disease management (Peeters et al., 2018; Peeters, 2017). The rationale being that, as disease activity and treatment response result from multiple interactions between various factors, classification of individuals into subpopulations with differences in prognosis or response to a specific treatment will require a vast amount of data (Peeters, 2017). Conceivably, the use of large data sets in combination with deep learning approaches may lead to evidence- and practice-based treatment algorithms (Butzkueven et al., 2006). To optimally realize the potential of existing and future data, the application of the Findable, Accessible, Interoperable and Reusable (FAIR) data concept has been proposed, as well as the development of multidisciplinary data infrastructures (Peeters, 2017; Wilkinson et al., 2016; Peeters et al.,

2018). In all, these concepts and their implementation will expectedly lead to sophisticated decision-support systems, and may also provide health insurers and regulators with long-term data on (relative) effectiveness and costs of treatments (Peeters, 2017).

4. Observational studies in pharmacotherapeutic research

Observational studies play a pivotal role in pharmacotherapeutic research: as reports of accidental observations, as phase 1 and 4 studies, and as alternatives to phase 3 RCTs.

4.1. Accidental observations

Incidentally, accidental observations of unexpected positive clinical changes during treatment with an authorized drug, in symptoms or signs other than those of the disease for which the drug is given, may lead to further research into the drug's efficacy in other conditions (Halvorsen and Martensen-Larsen, 1978). E.g. fumaric acid esters had been used successfully for the treatment of psoriasis vulgaris since 1959, when observations in patients who had both psoriasis and RRMS, suggested their efficacy in MS (Meissner et al., 2012; Liu et al., 2018). In general, accidental observations are valid methods for making discoveries regarding unknown effects of established treatments (Vandenbroucke, 2008).

4.2. Phase 1 studies

Phase 1 studies are uncontrolled trials in patients or normal persons to investigate pharmacokinetics, pharmacodynamics and safety aspects of different dosages of a drug. Often these are first-in-human studies, involving about ten to 20 persons (Streeter et al., 2015; Kanhai et al., 2018). Although patients may show improvement in surrogate or clinical outcomes, phase 1 studies cannot inform on the drug's efficacy, due to the small sample size and confounding variables (Sedel et al., 2015). E.g., following accidental observations in patients with psoriasis and RRMS, an exploratory, prospective, uncontrolled study in ten patients with RRMS and at least one gadolinium (Gd)-enhancing lesion on T1-weighted brain MRI, found that treatment with 360 mg/day and 720 mg/day of fumaric acid esters was associated with statistically significant reductions in the number and volume of Gd-enhancing lesions (Schimrigk et al., 2006).

4.3. Phase 4 studies

4.3.1. Effectiveness of treatment

In the post-authorization phase, studies of intended clinical changes during treatment with a drug whose efficacy has been demonstrated in RCTs, are crucial in pharmacotherapeutic research, as they establish the effectiveness of the treatment. Treatment effectiveness is not only determined by the drug's efficacy, but also by factors operative in real-life conditions, like reimbursement, social environment, treatment persistence and adherence to the dosing regimen. Most effectiveness studies are observational, especially ecological, cohort and case-control. As these studies generally include patients treated in daily practice, their findings can easily be applied to the overall MS population. On the other hand, if they are performed in academic hospitals or MS clinics, or in a single region, their generalizability is rather limited.

Three specific reasons for performing post-authorization effectiveness studies are mentioned with examples. First, the assessment of the relative effectiveness of two or more DMDs. E.g. to compare treatment effectiveness and persistence in RRMS patients who initiated rituximab vs. GA or INF-β, 461 patients from the Swedish MS registry in the rituximab arm were propensity score matched with 922 patients from the INF-β/GA comparator, showing that rituximab was associated with 87% reduction in ARR and 85% reduction in discontinuation rate relative to INF-β/GA (Spelman et al., 2018).

Second, the development of treatment strategies for patient subpopulations. The large number of DMD treatments in RRMS urges the establishment of optimal treatment sequences for different patient subpopulations. E.g. in patients who have to switch from a given DMD due to persistent disease activity vs. those who have to switch because of side effects or safety risks. Thus, Alping et al. compared outcomes for all RRMS patients switching from natalizumab due to JC virus antibody positivity at 3 Swedish MS centers, and findings suggested an improved effectiveness and tolerability of rituximab compared with fingolimod in stable patients (Alping et al., 2016). So, observational data contribute to a differentiated use of DMDs and the development of optimal treatment sequences in subpopulations of patients.

Third, the assessment of variation in the drug's efficacy. The results of large-scale international RCTs may show variation in beneficial effects among subgroups of patients, that are defined by country, geographic region or otherwise (Yusuf and Wittes, 2016). As these variations can occasionally reflect differences in true benefits, additional observational data – derived form treatment in daily practice – are needed to determine whether the RCT results are likely to be real or not (Yusuf and Wittes, 2016). However, variations in treatment effectiveness in real-life patients may also relate to pre-existing differences, e.g. in co-morbidities or co-medications. E.g. the lipid-lowering statins have anti-inflammatory effects that might be therapeutic in MS (Chan et al., 2017).

4.3.2. Validation of side effect profiles

In phase 2 and 3 RCTs selected patients are treated in settings not representative of daily practice. To increase the clinical usefulness of the RCTs' data on side effects, they have to be externally validated, especially in patients who would not have met the RCT's eligibility criteria, e.g. due to concomitant disease or age (Rothwell, 2005).

E.g., using quality assurance measures similar to those in phase 3 trials, an international prospective cohort study found that 75% of intramuscular INF-β-1a-treated RRMS patients experienced one or more adverse events that were likely or definitely related to INF-β-1a, and could therefore be qualified as side effects, 69% of these being mild (Jongen et al., 2011). Moreover, in comparison with the pivotal intramuscular INF-β phase 3 trial, statistically significantly lower incidences were found for most of the common side effects, which suggests that in daily practice clinically relevant – i.e. reported after explicit questioning – side effects are less frequent than indicated by phase 3 RCT data (Jongen et al., 2011).

Notably, in phase 2 and 3 RCTs side effects are systematically evaluated by researchers asking specific predefined questions, whereas physicians in daily practice may have to confine themselves to a single open question. Moreover, more than a few daily-practice studies on side effects are retrospective, single-centered, restricted to academic settings, or limited to one region or country (Barra et al., 2016). Ideally, side effect validation studies are performed prospectively in general and academic practices, in different countries and regions, and – to enhance data quality – apply the International Conference on Harmonization (ICH) and Good Clinical Practice (GCP) guidelines, supported by a clinical research organization (CRO) (Jongen et al., 2011; Jongen, 2017b).

4.3.3. Detection of rare or long-term side effects

Given their limited duration, restricted numbers of patients and the homogeneity of patient groups, phase 2 and 3 RCTs conceivably cannot generate adequate side effect profiles. In contrast, observational phase 4 studies are well suited to identify and quantify side effects, especially those that are rare or occur over the longer term (de Jong et al., 2017). Nevertheless, confounding by indication may be a problem in these studies, especially when assessing the incidence of side effects and when comparing incidences between treatments (Vandenbroucke, 2008). E.g. in clinical practice, risk factors for migraine or stroke might be linked to the choice of treatment (age, sex,

body mass index, socioeconomic factors, preexisting co-morbidities, mobility) and may thus cause different rates of these outcomes in patients with different treatments.

Unfortunately, most post-authorization studies in MS insufficiently assess side effects. Thus, it was two decades after the introduction of the first INF-β drug before an increased risk of stroke, migraine, depression and hematologic abnormalities was convincingly demonstrated in INF-β-treated patients (de Jong et al., 2017). Large-scale, prospective, ICH-GCP-based and CRO-supported observational studies, initiated immediately after market authorization, are optimally suited to the timely detection of rare or long-term side effects and safety signals (Jongen et al., 2011; Jongen, 2017b). Special attention should be given to biosimilars and follow-on non-biological complex drugs, as their safety data are limited at the time of market authorization due to relatively small numbers of patient-years .

4.4. Alternative to randomized controlled trials

When ethical, practical or financial considerations prevent large RCTs from being conducted, observational studies may provide second-best evidence for the beneficial effects of treatment. E.g. it is difficult to obtain funding for RCTs when drugs are off-label or patents have expired. In these and similar conditions observational studies may be a realistic option to acquire actionable data (Frieden, 2017).

5. Discussion

Observational studies play a major role in clinical MS research. Whether observational findings can be qualified as demonstrative, indicative or only suggestive, mainly depends on four factors: the research question, whether and how the design fits this question, the analytical techniques applied, and the quality of the data. Thus, when assessing intended treatment effects, RCT follow-up studies provide stronger evidence than case series, whereas, when looking for discoveries, the evidence of case series may be as strong as that of case-control studies. The appropriateness of analytical techniques relates to the research question and study design, and as a review of these aspects is not within the scope of this article the reader is referred to other publications (Kalincik and Butzkueven, 2016; Sormani and Bruzzi, 2015; Trojano et al., 2009). Several questions arise when assessing the evidential value of observational studies with respect to data quality, such as: was the study prospective, retrospective or cross-sectional; who generated the data, e.g. who performed EDSS scoring; what measures were taken to minimise inter-rater or inter-centre differences, e.g. via standardization of EDSS scoring (Kappos et al., 2015a), MRI settings and MRI quantification methods (Giorgio and De Stefano, 2018); who did the data input; and which measures were taken to assure data quality. These issues pertain a fortiori to very large-scale and multi-disciplinary longitudinal studies – often registry-based – given the sheer quantity of data and the diversity of data sources (Insel, 2019).

At the stage of drug approval, regulatory authorities have to rely principally on phase 2 and 3 RCT data to evaluate adverse events and safety risks (Schultz, 2007). After market authorization, (registry-based) cohort studies may extensively investigate side effects and safety issues in the real world, including previously unrecognized concerns (Ording et al., 2016; Services, 2007). Actually, the European Union pharmacovigilance system aims to continuously obtain information on the safety of drugs, through spontaneous case reports, risk management programs and observational studies (Santoro et al., 2017). National registries may also focus on safety and side effects;. e.g. as part of the Swedish MS registry, a national post-marketing surveillance program monitors the safety of natalizumab (Piehl et al., 2011).

The role of observational data in assessing the effectiveness of treatments is often disputed, given the potential influence of confounding factors. It has even been suggested that results from observational studies on the effectiveness of DMD treatments in MS should not be relied on for clinical or public health decisions (Sormani and Bruzzi, 2015). On the other hand, for many clinical and health policy questions, the only relevant information available to aid in decision making is based on observational data (Van Poucke et al., 2016). Actually, observational evidence for decisive clinical action is receiving increased interest, as are ways to improve the use of multiple data sources – including observational ones – for making treatment decisions (Frieden, 2017; Peeters, 2017; Peeters et al., 2018). As to DMD treatment in MS, the combined evidence from a large number of (registry-based) cohort, RCT follow-up and ecological studies sufficiently convinced neurologists and health authorities of the effectiveness of these treatments to continue using them. This also holds for the view that, in general, early DMD treatment is better than late treatment, and that late treatment is better than no treatment at all (Cocco et al., 2015). Actually, as observational DMD studies are typically situated in real life settings, they are particularly suited to obtaining information that is relevant to neurologists – e.g. regarding stabilisation or improvement of disability (EDSS score) or 'no evidence of disease activity' (NEDA) status (Parks et al., 2017) and patients - e.g. regarding stabilisation or improvement of cumbersome symptoms (e.g. fatigue) or overall condition (HRQoL) (Jongen, 2017a), when they have to decide on treatment options (Jongen et al., 2016). In all, obtaining actionable data from observational research is a reasonable and achievable goal (Frieden, 2017). Thus, by using various designs and settings, a body of evidence is now being built on the effectiveness of specific DMD treatments and their optimal sequences in subgroups of MS patients (Pardo and Jones, 2017). Yet, given the very nature of their design, observational studies may also achieve contrary or contradictory results. E.g. meta-analyses of high quality case–control studies suggested a statistically significant association between pre-morbid head trauma and the risk for developing MS, whereas cohort studies did not (Lunny et al., 2014).

A problem that arises when analyzing pooled data from various data bases, is the use of different measures of the same clinical phenomenon. E.g. for the assessment of MS-related fatigue at least five validated scales are being used in clinical practice and research; and for HRQoL the common use of at least seven scales has been reported (Jongen, 2017a). As it seems unfeasible or undesirable to try to agree worldwide upon a single scale for a given symptom or sign, consensus should be reached on a limited set of measures and methods should be developed that convert various measures (Peeters et al., 2018).

Observational studies are increasingly embedded in registries. National quality registries use a multi-stakeholder approach that benefits governmental health organisations by supplying data about the use of health services, including DMDs (Fagius et al., 2017; Johansson et al., 2009); healthcare providers by presenting a data overview and follow-up charts (Hillert and Stawiarz, 2015); and patients by providing a summary of personal data and the opportunity to generate PROs (Hillert and Stawiarz, 2015; Butzkueven et al., 2006). By combining physician-reported, paraclinical and patient-reported outcomes, and by linkage to other data sets, MS registries become platforms for integrated research (Hillert and Stawiarz, 2015). However, for very large-scale registry-based studies, especially international ones, the conclusions' relevance for individual patients is not self-evident. E.g. treatment effectiveness partly depends on factors that may vary by country or region, and that are usually not documented, like care settings, socio-economic circumstances and patient-caregiver relationship (Insel, 2019; Harding et al., 2019).

Well-defined criteria for reporting observational effectiveness studies – similar to the CONSORT guidelines for RCT reporting (Moher et al., 2001) – are lacking, as the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines are devised for epidemiological research (von Elm et al., 2007). Consequently, observational reports on treatment effects may fail to provide essential information that is necessary for understanding the quality of the study and the relevance of its conclusions (French and Gronseth, 2008;

Barnish and Turner, 2017). Therefore, we suggest that reports of observational effectiveness studies describe in the abstract the following six items: question(s) the study meant to answer, study design, patient population, treatment of interest, relevant outcomes, and strengths and limitations of the study. To further improve the quality of reporting and to prevent incomplete or selective reporting, an extension of STROBE may be considered: Strengthening the Reporting of Observational Studies in Effectiveness Research (STROBER) guidelines (von Elm et al., 2007). Furthermore, to increase among neurologists the knowledge about methodological and quality aspects of observational studies, these topics need to be given more attention in scientific MS journals and at MS congresses (Winchester, 2018; Saposnik, 2019).

Future directions in observational MS research may comprise the systematic inclusion and valuation of patient-reported and economic outcomes (Kalincik and Butzkueven, 2016), the application of new analytical methods, the development of predictive and treatment algorithms (Trojano et al., 2017), and - perhaps most important - the use of pharmacovigilance systems that guarantee the very early detection of unknown, possibly life-threatening side effects (Muraro et al., 2018).

Funding

I did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors for the generation of this manuscript, nor any other funding.

Conflict of interest

I have received honoraria from Bayer Netherlands for consultancy activities.

References

Alping, P., Frisell, T., Novakova, L., Islam-Jakobsson, P., Salzer, J., Bjorck, A., Axelsson, M., Malmestrom, C., Fink, K., Lycke, J., Svenningsson, A., Piehl, F., 2016. Rituximab versus fingolimod after natalizumab in multiple sclerosis patients. Ann. Neurol. 79 (6), 950–958.

Alwan, S., Chambers, C.D., Armenti, V.T., Sadovnick, A.D., 2015. The need for a disease-specific prospective pregnancy registry for multiple sclerosis (MS). Mult. Scler. Relat. Disord. 4 (1), 6–17.

Andersen, O., 2012. From the Gothenburg cohort to the Swedish multiple sclerosis registry. Acta Neurol. Scand. (Supplementum(195)), S13–S19.

Arnal-Garcia, C., Amigo-Jorrin Mdel, C., Lopez-Real, A.M., Lema-Devesa, C., Llopis, N., Sanchez-de la Rosa, R., Group, X.-S., 2014. Long-term effectiveness of glatiramer acetate in clinical practice conditions. J. Clin. Neurosci. 21 (12), 2212–2218.

Bailar III JC, M.F., 1986. Medical uses of statistics. The Massachusetts Medical Society. Waltham, Massachusetts.

Barnish, M.S., Turner, S., 2017. The value of pragmatic and observational studies in health care and public health. Pragmat. Obs. Res. 8, 49–55.

Barra, M.E., Soni, D., Vo, K.H., Chitnis, T., Stankiewicz, J.M., 2016. Experience with long-term rituximab use in a multiple sclerosis clinic. Mult. Scler. J. – Exp. Transl. Clin. 2, 2055217316672100.

Boesen, M.S., Koch-Henriksen, N., Thygesen, L.C., Eriksson, F., Greisen, G., Born, A.P., Blinkenberg, M., Uldall, P.V., Magyari, M., 2018. Infections seem to be more frequent before onset of pediatric multiple sclerosis: a Danish nationwide nested case-control study. Mult. Scler., 1352458518771871.

Brown, J.W.L., Coles, A., Horakova, D., Havrdova, E., Izquierdo, G., Prat, A., Girard, M., Duquette, P., Trojano, M., Lugaresi, A., Bergamaschi, R., Grammond, P., Alroughani, R., Hupperts, R., McCombe, P., Van Pesch, V., Sola, P., Ferraro, D., Grand'Maison, F., Terzi, M., Lechner-Scott, J., Flechter, S., Slee, M., Shaygannejad, V., Pucci, E., Granella, F., Jokubaitis, V., Willis, M., Rice, C., Scolding, N., Wilkins, A., Pearson, O.R., Ziemssen, T., Hutchinson, M., Harding, K., Jones, J., McGuigan, C., Butzkueven, H., Kalincik, T., Robertson, N., Group, M.S.S., 2019. Association of initial disease-modifying therapy with later conversion to secondary progressive multiple sclerosis. JAMA 321 (2), 175–187.

Butzkueven, H., Chapman, J., Cristiano, E., Grand'Maison, F., Hoffmann, M., Izquierdo, G., Jolley, D., Kappos, L., Leist, T., Pohlau, D., Rivera, V., Trojano, M., Verheul, F., Malkowski, J.P., 2006. MSBase: an international, online registry and platform for collaborative outcomes research in multiple sclerosis. Mult. Scler. 12 (6), 769–774.

Cadden, M.H., Arnett, P.A., Tyry, T.M., Cook, J.E., 2018. Judgment hurts: the psychological consequences of experiencing stigma in multiple sclerosis. Soc. Sci. Med.

Chan, D., Binks, S., Nicholas, J.M., Frost, C., Cardoso, M.J., Ourselin, S., Wilkie, D., Nicholas, R., Chataway, J., 2017. Effect of high-dose simvastatin on cognitive, neuropsychiatric, and health-related quality-of-life measures in secondary progressive multiple sclerosis: secondary analyses from the MS-STAT randomised, placebo-controlled trial. Lancet Neurol. 16 (8), 591–600.

Cocco, E., Sardu, C., Spinicci, G., Musu, L., Massa, R., Frau, J., Lorefice, L., Fenu, G., Coghe, G., Massole, S., Maioli, M.A., Piras, R., Melis, M., Porcu, G., Mamusa, E., Carboni, N., Contu, P., Marrosu, M.G., 2015. Influence of treatments in multiple sclerosis disability: a cohort study. Mult. Scler. 21 (4), 433–441.

de Jong, H.J.I., Kingwell, E., Shirani, A., Cohen Tervaert, J.W., Hupperts, R., Zhao, Y., Zhu, F., Evans, C., van der Kop, M.L., Traboulsee, A., Gustafson, P., Petkau, J., Marrie, R.A., Tremlett, H., British Columbia Multiple Sclerosis Clinic, N., 2017. Evaluating the safety of beta-interferons in MS: a series of nested case-control studies. Neurology.

Essebag, V., Genest Jr., J., Suissa, S., Pilote, L., 2003. The nested case-control study in cardiology. Am. Heart J. 146 (4), 581–590.

Fagius, J., Feresiadou, A., Larsson, E.M., Burman, J., 2017. Discontinuation of disease modifying treatments in middle aged multiple sclerosis patients. first line drugs vs natalizumab. Mult. Scler. Relat. Disord. 12, 82–87.

Feigin, V.L., Howard, G., 2008. The importance of epidemiological studies should not be downplayed. Stroke J. Cereb. Circ. 39 (1), 1–2.

Fitzgerald, K.C., Tyry, T., Salter, A., Cofield, S.S., Cutter, G., Fox, R.J., Marrie, R.A., 2018. A survey of dietary characteristics in a large population of people with multiple sclerosis. Mult. Scler. Relat. Disord. 22, 12–18.

Flachenecker, P., Buckow, K., Pugliatti, M., Kes, V.B., Battaglia, M.A., Boyko, A., Confavreux, C., Ellenberger, D., Eskic, D., Ford, D., Friede, T., Fuge, J., Glaser, A., Hillert, J., Holloway, E., Ioannidou, E., Kappos, L., Kasilingam, E., Koch-Henriksen, N., Kuhle, J., Lepore, V., Middleton, R., Myhr, K.M., Orologas, A., Otero, S., Pitschnau-Michel, D., Rienhoff, O., Sastre-Garriga, J., Schyns-Liharska, T., Sutovic, D., Thalheim, C., Trojano, M., Vlasov, Y.V., Yaldizli, O., Consortium, E.U., 2014. Multiple sclerosis registries in Europe – results of a systematic survey. Mult. Scler. 20 (11), 1523–1532.

Ford, C., Goodman, A.D., Johnson, K., Kachuck, N., Lindsey, J.W., Lisak, R., Luzzio, C., Myers, L., Panitch, H., Preiningerova, J., Pruitt, A., Rose, J., Rus, H., Wolinsky, J., 2010. Continuous long-term immunomodulatory therapy in relapsing multiple sclerosis: results from the 15-year analysis of the US prospective open-label study of glatiramer acetate. Mult. Scler. 16 (3), 342–350.

Ford, C.C., Johnson, K.P., Lisak, R.P., Panitch, H.S., Shifronis, G., Wolinsky, J.S., Copaxone Study, G., 2006. A prospective open-label study of glatiramer acetate: over a decade of continuous use in multiple sclerosis patients. Mult. Scler. 12 (3), 309–320.

French, J., Gronseth, G., 2008. Lost in a jungle of evidence: we need a compass. Neurology 71 (20), 1634–1638.

Frieden, T.R., 2017. Evidence for health decision making – beyond randomized, controlled trials. N. Engl. J. Med. 377 (5), 465–475.

Ghezzi, A., Moiola, L., Pozzilli, C., Brescia-Morra, V., Gallo, P., Grimaldi, L.M., Filippi, M., G, G.C., Neurology, M.S.S.G.-I.S.o., 2015. Natalizumab in the pediatric MS population: results of the Italian registry. BMC Neurol. 15, 174.

Giorgio, A., De Stefano, N., 2018. Effective utilization of MRI in the diagnosis and management of multiple sclerosis. Neurol. Clin. 36 (1), 27–34.

Glasziou, P., Chalmers, I., Rawlins, M., McCulloch, P., 2007. When are randomised trials unnecessary? Picking signal from noise. BMJ 334 (7589), 349–351.

Granqvist, M., Boremalm, M., Poorghobad, A., Svenningsson, A., Salzer, J., Frisell, T., Piehl, F., 2018. Comparative effectiveness of rituximab and other initial treatment choices for multiple sclerosis. JAMA Neurol.

Halvorsen, K.A., Martensen-Larsen, O., 1978. Apomorphine revived: fortified, prolonged, and improved therapeutical effect. Int. Jo. Addict. 13 (3), 475–484.

Harding, K.E., Wardle, M., Carruthers, R., Robertson, N., Zhu, F., Kingwell, E., Tremlett, H., 2019. Socioeconomic status and disability progression in multiple sclerosis: a multinational study. Neurology.

Hatam, N., Bastani, P., Shahtaheri, R.S., 2016. Quality of life in relapsing-remitting multiple sclerosis patients receiving CinnoVex compared with Avonex. J. Res. Pharmacy Pract. 5 (3), 181–185.

Hernan, M.A., Hernandez-Diaz, S., Robins, J.M., 2004. A structural approach to selection bias. Epidemiology 15 (5), 615–625.

Hillert, J., Stawiarz, L., 2015. The Swedish MS registry – clinical support tool and scientific resource. Acta Neurol. Scand. 132 (Supplementum (199)), S11–S19.

Insel, T.R., 2019. How algorithms could bring empathy back to medicine. Nature 567 (7747), 172–173.

Johansson, S., Ytterberg, C., Gottberg, K., Widen Holmqvist, L., von Koch, L., 2009. Use of health services in people with multiple sclerosis with and without fatigue. Mult. Scler. 15 (1), 88–95.

Jongen, P.J., 2017a. Health-related quality of life in patients with multiple sclerosis: impact of disease-modifying drugs. CNS Drugs 31 (7), 585–602.

Jongen, P.J., 2017b. Letter re: evaluating the safety of beta-interferons in MS: a series of nested case-control studies. Neurology 89 (19), 2022.

Jongen, P.J., Lehnick, D., Koeman, J., Frequin, S., Heersema, D., Kornips, B., Schyns-Soeterboek, A., Visser, L.H., Schiphof, P., Valkenburg, A., Hiel, J., 2014. Fatigue and health-related quality of life in relapsing-remitting multiple sclerosis after 2 years glatiramer acetate treatment are predicted by changes at 6 months: an observational multi-center study. J. Neurol.

Jongen, P.J., Lehnick, D., Sanders, E., Seeldrayers, P., Fredrikson, S., Andersson, M., Speck, J., 2010a. Health-related quality of life in relapsing remitting multiple sclerosis patients during treatment with glatiramer acetate: a prospective, observational, international, multi-centre study. Health Qual. Life Outcomes 8, 133.

Jongen, P.J., Lemmens, W.A., Hupperts, R., Hoogervorst, E., Schrijver, H.M., Slettenaar, A., de Schryver, E.L., Boringa, J., van Noort, E., Donders, R., 2016. Persistence and adherence in multiple sclerosis patients starting glatiramer acetate treatment: assessment of relationship with care received from multiple disciplines. Patient Prefer Adherence 10, 909–917.

Jongen, P.J., Sindic, C., Carton, H., Zwanikken, C., Lemmens, W., Borm, G., 2010b.

Improvement of health-related quality of life in relapsing remitting multiple sclerosis patients after 2 years of treatment with intramuscular interferon-beta-1a. J. Neurol. 257 (4), 584–589.
Jongen, P.J., Sindic, C., Sanders, E., Hawkins, S., Linssen, W., van Munster, E., Frequin, S., Borm, G., Functional, C., Quality of Life in Avonex-treated Relapsing Multiple Sclerosis Patients Study, G., 2011. Adverse events of interferon beta-1a: a prospective multi-centre international ICH-GCP-based CRO-supported external validation study in daily practice. PLoS One 6 (10), e26568.
Kalincik, T., Butzkueven, H., 2016. Observational data: understanding the real MS world. Mult. Scler. 22 (13), 1642–1648.
Kanhai, K.M.S., Zuiker, R., Stavrakaki, I., Gladdines, W., Gaillard, P.J., Klaassen, E.S., Groeneveld, G.J., 2018. Glutathione-PEGylated liposomal methylprednisolone in comparison to free methylprednisolone: slow release characteristics and prolonged lymphocyte depression in a first-in-human study. Br. J. Clin. Pharmacol. 84 (5), 1020–1028.
Kappos, L., D'Souza, M., Lechner-Scott, J., Lienert, C., 2015a. On the origin of Neurostatus. Mult. Scler. Relat. Disord. 4 (3), 182–185.
Kappos, L., Kuhle, J., Multanen, J., Kremenchutzky, M., Verdun di Cantogno, E., Cornelisse, P., Lehr, L., Casset-Semanaz, F., Issard, D., Uitdehaag, B.M., 2015b. Factors influencing long-term outcomes in relapsing-remitting multiple sclerosis: PRISMS-15. J. Neurol. Neurosurg. Psychiatry 86 (11), 1202–1207.
Kappos, L., Traboulsee, A., Constantinescu, C., Eralinna, J.P., Forrestal, F., Jongen, P., Pollard, J., Sandberg-Wollheim, M., Sindic, C., Stubinski, B., Uitdehaag, B., Li, D., 2006. Long-term subcutaneous interferon beta-1a therapy in patients with relapsing-remitting MS. Neurology 67 (6), 944–953.
Kleinschmidt-DeMasters, B.K., Tyler, K.L., 2005. Progressive multifocal leukoencephalopathy complicating treatment with natalizumab and interferon beta-1a for multiple sclerosis. N. Engl. J. Med. 353 (4), 369–374.
Koch-Henriksen, N., Magyari, M., Laursen, B., 2015. Registers of multiple sclerosis in Denmark. Acta Neurol. Scand. 132 (199), 4–10.
Kyriacou, D.N., Lewis, R.J., 2016. Confounding by indication in clinical research. JAMA 316 (17), 1818–1819.
Landfeldt, E., Castelo-Branco, A., Svedbom, A., Lofroth, E., Kavaliunas, A., Hillert, J., 2018. The long-term impact of early treatment of multiple sclerosis on the risk of disability pension. J. Neurol. 265 (3), 701–707.
Langer-Gould, A., Atlas, S.W., Green, A.J., Bollen, A.W., Pelletier, D., 2005. Progressive multifocal leukoencephalopathy in a patient treated with natalizumab. N. Engl. J. Med. 353 (4), 375–381.
Liu, C.Y., Tung, T.H., Lee, C.Y., Chang, K.H., Wang, S.H., Chi, C.C., 2018. Association of multiple sclerosis with psoriasis: a systematic review and meta-analysis of observational studies. Am. J. Clin. Dermatol.
Lunny, C.A., Fraser, S.N., Knopp-Sihota, J.A., 2014. Physical trauma and risk of multiple sclerosis: a systematic review and meta-analysis of observational studies. J. Neurol. Sci. 336 (1–2), 13–23.
Magyari, M., Koch-Henriksen, N., Sorensen, P.S., 2016. The Danish multiple sclerosis treatment register. Clin. Epidemiol. 8, 549–552.
Mansournia, M.A., Higgins, J.P., Sterne, J.A., Hernan, M.A., 2017. Biases in randomized trials: a conversation between trialists and epidemiologists. Epidemiology 28 (1), 54–59.
Meissner, M., Valesky, E.M., Kippenberger, S., Kaufmann, R., 2012. Dimethyl fumarate – only an anti-psoriatic medication? J. Dtsh. Dermatol. Ges. 10 (11), 793–801.
Merriam-Webster, 2018.
Moher, D., Schulz, K.F., Altman, D., Group, C., 2001. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. JAMA 285 (15), 1987–1991.
Moisset, X., Perie, M., Pereira, B., Dumont, E., Lebrun-Frenay, C., Lesage, F.X., Dutheil, F., Taithe, F., Clavelou, P., 2017. Decreased prevalence of cancer in patients with multiple sclerosis: a case-control study. PLoS One 12 (11), e0188120.
Muraro, P.A., Scolding, N.J., Fox, R.J., 2018. Rare side effects of alemtuzumab remind us of the need for postmarketing surveillance. Neurology 90 (18), 819–820.
Ording, A.G., Cronin-Fenton, D., Ehrenstein, V., Lash, T.L., Acquavella, J., Rorth, M., Sorensen, H.T., 2016. Challenges in translating endpoints from trials to observational cohort studies in oncology. Clin. Epidemiol. 8, 195–200.
Palace, J., Duddy, M., Bregenzer, T., Lawton, M., Zhu, F., Boggild, M., Piske, B., Robertson, N.P., Oger, J., Tremlett, H., Tilling, K., Ben-Shlomo, Y., Dobson, C., 2015. Effectiveness and cost-effectiveness of interferon beta and glatiramer acetate in the UK Multiple Sclerosis Risk Sharing Scheme at 6 years: a clinical cohort study with natural history comparator. Lancet Neurol. 14 (5), 497–505.
Pardo, G., Jones, D.E., 2017. The sequence of disease-modifying therapies in relapsing multiple sclerosis: safety and immunologic considerations. J. Neurol. 264 (12), 2351–2374.
Parks, N.E., Flanagan, E.P., Lucchinetti, C.F., Wingerchuk, D.M., 2017. NEDA treatment target? No evident disease activity as an actionable outcome in practice. J. Neurol. Sci. 383, 31–34.
Peeters, L.M., 2017. Fair data for next-generation management of multiple sclerosis. Mult. Scler., 1352458517748475.
Peeters, L.M., van Munster, C.E., Van Wijmeersch, B., Bruyndonckx, R., Lamers, I., Hellings, N., Popescu, V., Thalheim, C., Feys, P., 2018. Multidisciplinary data infrastructures in multiple sclerosis: why they are needed and can be done!. Mult. Scler., 1352458518807076.
Piehl, F., Holmen, C., Hillert, J., Olsson, T., 2011. Swedish natalizumab (Tysabri) multiple sclerosis surveillance study. Neurol. Sci. 31 (Suppl 3), S289–S293.
Prosperini, L., de Rossi, Scarpazza, N., Moiola, C., Cosottini, L., Gerevini, M., Capra, S., Italian, R., s.g., P.M.L., 2016. Natalizumab-related progressive multifocal leukoencephalopathy in multiple sclerosis: findings from an italian independent registry. PLoS One 11 (12), e0168376.
Rothman KJ, G.S., Lash, T. (Eds.), 2008. Modern Epidemiology, Third ed. Lippincott Williams & Wilkins, Philadelphia, PA.
Rothwell, P.M., 2005. External validity of randomised controlled trials: "to whom do the results of this trial apply? Lancet 365 (9453), 82–93.
Salter, A.R., Marrie, R.A., Agashivala, N., Belletti, D.A., Kim, E., Cutter, G.R., Cofield, S.S., Tyry, T., 2014. Patient perspectives on switching disease-modifying therapies in the NARCOMS registry. Patient Prefer Adher. 8, 971–979.
Santoro, A., Genov, G., Spooner, A., Raine, J., Arlett, P., 2017. Promoting and protecting public health: how the European union pharmacovigilance system works. Drug Saf. 40 (10), 855–869.
Saposnik, G., Maurino, J., Sempere, A.P., Terzaghi, M.A., Amato, M.P., Montalban, X., 2019. Does attendance at the ECTRIMS congress impact on therapeutic decisions in multiple sclerosis care? Mult. Scler. J. 1–5.
Schimrigk, S., Brune, N., Hellwig, K., Lukas, C., Bellenberg, B., Rieks, M., Hoffmann, V., Pohlau, D., Przuntek, H., 2006. Oral fumaric acid esters for the treatment of active multiple sclerosis: an open-label, baseline-controlled pilot study. Eur. J. Neurol. 13 (6), 604–610.
Schultz, W.B., 2007. Bolstering the FDA's drug-safety authority. N. Engl. J. Med. 357 (22), 2217–2219.
Sedel, F., Papeix, C., Bellanger, A., Touitou, V., Lebrun-Frenay, C., Galanaud, D., Gout, O., Lyon-Caen, O., Tourbah, A., 2015. High doses of biotin in chronic progressive multiple sclerosis: a pilot study. Mult. Scler. Relat. Disord. 4 (2), 159–169.
Services, U.S.D.o.H.a.H, 2007. The Future of Drug Safety - Promoting and Protecting the Health of the Public: FDA's Response to the Institute of Medicine's 2006 Report.
Smith, G.C., Pell, J.P., 2003. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. BMJ 327 (7429), 1459–1461.
Sormani, M.P., Bruzzi, P., 2015. Can we measure long-term treatment effects in multiple sclerosis? Nat. Rev. Neurol. 11 (3), 176–182.
Spelman, T., Frisell, T., Piehl, F., Hillert, J., 2018. Comparative effectiveness of rituximab relative to INF-beta or glatiramer acetate in relapsing-remitting MS from the Swedish MS registry. Mult. Scler. 24 (8), 1087–1095.
Streeter, H.B., Rigden, R., Martin, K.F., Scolding, N.J., Wraith, D.C., 2015. Preclinical development and first-in-human study of ATX-MS-1467 for immunotherapy of MS. Neurol. Neuroimmunol. Neuroinflamm. 2 (3), e93.
Thrift, A., 2009. Types of epidemiological studies and their uses: with a focus on population surveys. In: International Conference on Clinical Neuroepidemiology, Munich.
Trojano, M., Pellegrini, F., Fuiani, A., Paolicelli, D., Zipoli, V., Zimatore, G.B., Di Monte, E., Portaccio, E., Lepore, V., Livrea, P., Amato, M.P., 2007. New natural history of interferon-beta-treated relapsing multiple sclerosis. Ann. Neurol. 61 (4), 300–306.
Trojano, M., Pellegrini, F., Paolicelli, D., Fuiani, A., Di Renzo, V., 2009. Observational studies: propensity score analysis of non-randomized data. Int. MS J. 16 (3), 90–97.
Trojano, M., Tintore, M., Montalban, X., Hillert, J., Kalincik, T., Iaffaldano, P., Spelman, T., Sormani, M.P., Butzkueven, H., 2017. Treatment decisions in multiple sclerosis – insights from real-world observational studies. Nat. Rev. Neurol. 13 (2), 105–118.
Van Assche, G., Van Ranst, M., Sciot, R., Dubois, B., Vermeire, S., Noman, M., Verbeeck, J., Geboes, K., Robberecht, W., Rutgeerts, P., 2005. Progressive multifocal leukoencephalopathy after natalizumab therapy for Crohn's disease. N. Engl. J. Med. 353 (4), 362–368.
Van Poucke, S., Thomeer, M., Heath, J., Vukicevic, M., 2016. Are randomized controlled trials the (G)old standard? From clinical intelligence to prescriptive analytics. J. Med. Internet Res. 18 (7), e185.
Vandenbroucke, J.P., 2008. Observational research, randomised trials, and two views of medical science. PLoS Med. 5 (3), e67.
Voldsgaard, A., Koch-Henriksen, N., Magyari, M., Sellebjerg, F., Sorensen, P.S., Oturai, A.B., 2017. Early safety and efficacy of fingolimod treatment in Denmark. Acta Neurol. Scand. 135 (1), 129–133.
von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gotzsche, P.C., Vandenbroucke, J.P., 2007. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 335 (7624), 806–808.
Wang, G., Marrie, R.A., Fox, R.J., Tyry, T., Cofield, S.S., Cutter, G.R., Salter, A., 2018. Treatment satisfaction and bothersome bladder, bowel, sexual symptoms in multiple sclerosis. Mult. Scler. Relat. Disord. 20, 16–21.
Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J., Groth, P., Goble, C., Grethe, J.S., Heringa, J., t Hoen, P.A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B., 2016. The FAIR guiding principles for scientific data management and stewardship. Sci. Data 3, 160018.
Winchester, C., 2018. Give every paper a read for reproducibility. Nature 557 (7705), 281.
Yusuf, S., Wittes, J., 2016. Interpreting geographic variations in results of randomized, controlled trials. N. Engl. J. Med. 375 (23), 2263–2271.