BMC Medicine

OPINION   Open Access

The importance of collecting structured clinical information on multiple sclerosis

Tjalf Ziemssen[1*], Jan Hillert[2] and Helmut Butzkueven[3]

Abstract

Background: Randomized controlled trials (RCTs) are the 'gold standard' in the generation of drug efficacy and safety evidence. However, enrolment criteria, timelines and atypical comparators of RCTs limit their relevance to standard clinical practice.

Discussion: Real-world data (RWD) provide longitudinal information on the comparative effectiveness and tolerability of drugs, as well as their impact on resource use, medical costs, and pharmacoeconomic and patient-reported outcomes. This is particularly important in multiple sclerosis (MS), where economic treatment benefits of long-term disability reduction are a cornerstone of payer drug approvals – these are typically not examined in the RCT itself but modelled using real-world datasets. Importantly, surrogate markers used in RCTs to predict the prevention of long-term disability progression can only truly be assessed through RWD methodologies.

Summary: We discuss the differences between RCTs and RWD studies, describe how RWD complements the evidence base from RCTs in MS, summarize the different methods of RWD collection, and explain the importance of structuring data analysis to avoid bias. Guidance on performing and identifying high-quality real-world evidence studies is also provided.

Keywords: Multiple sclerosis, Real-world evidence, Real-world data, Randomised controlled trials, Registries, Pharmacoeconomics

An introduction to real-world evidence (RWE) in multiple sclerosis (MS)

In general, randomised controlled trials (RCTs) and RWE studies are important for improving our understanding of disease outcomes and treatment effects, with the two methodologies being increasingly viewed as complementary by clinicians, the pharmaceutical industry, drug regulatory and reimbursement agencies, and patients [1]. The prominence and value of RWE studies (Box 1) have made them mandatory in many settings; however, inherent variability in patient care means that their analysis requires particular care. Figure 1 summarises the key differences between RCTs and RWE studies.

In MS, there is a current and growing emphasis on obtaining data beyond phase 3 RCTs. In 2014, the number of published RWE studies in MS exceeded that of published phase 2 and 3 studies by more than two-fold (Fig. 2). This has been driven by increasing demand from payers and healthcare decision-makers for post-approval evidence to inform reviews of pricing, reimbursement, licences for new therapies, and formulation and indication changes [2]. Post-approval RWE studies are important for providing information on compliance with current treatment guidelines, identifying suboptimal therapies, defining treatment responder subgroups, optimizing treatment sequencing, and monitoring rare serious adverse events. This information can support licence extensions and treatment sequencing [3].

RWE studies can also provide valuable insights prior to product development. Pre-launch RWE studies, for example, are useful for mapping out the natural and drug-modified history of disease, current practice patterns and service structures [3]. In MS, this can go some way to meeting the need for information on disease characteristics, treatment behaviours, and healthcare availability for 'real' patients with the disease. Closer to

* Correspondence: tjalf.ziemssen@uniklinikum-dresden.de
[1]Center of Clinical Neuroscience, Department of Neurology, MS Center Dresden, Center of Clinical Neuroscience, University Hospital Carl Gustav Carus, Dresden University of Technology, Fetscherstr. 74, 01307 Dresden, Germany
Full list of author information is available at the end of the article

Ziemssen *et al. BMC Medicine* (2016) 14:81

Page 2 of 9

Box 1. About real-world evidence (RWE) studies [1, 5]

Randomised controlled trials (RCTs) are the 'gold standard' in the generation of efficacy and safety evidence of a product in restricted trial settings.
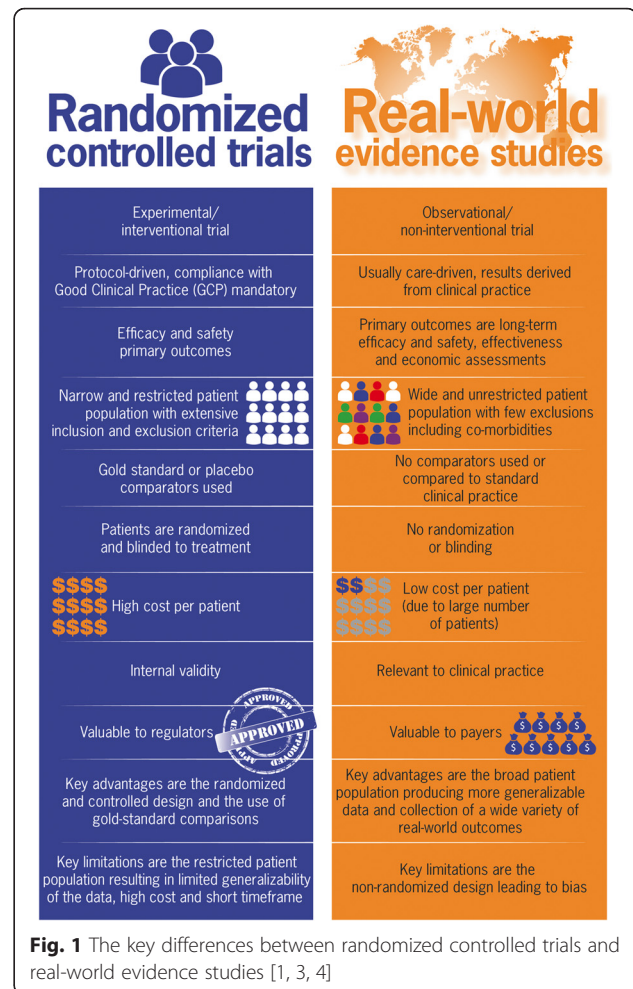
Advantages

1. Can mimic RCTs in the real world to allow assessment of a drug in the clinical setting

2. Complement the evidence base from RCTs by assessing a diverse range of outcome measures to provide information that may not be captured by other means, including:
   - comparative effectiveness data between multiple therapies
   - information on long-term disability outcomes
   - better characterization of long-term exposure risks and benefit–risk profiles
   - patient-reported outcomes
   - economic outcomes

Challenges

1. They cannot address the potential benefits and risks of a product in the real world [1, 5] because enrolment is restricted by disease-activity criteria, and subjects with comorbid conditions are typically excluded

2. Active comparator arms may not reflect the usual standard of care and selected endpoints can be artificial or chosen to maximize statistical power, thereby limiting the real-world relevance of study conclusions

3. There are significant ethical concerns about conducting placebo-controlled trials in countries where disease-modifying therapies are approved and reimbursed [2]; these pivotal trials are therefore conducted in countries without these approvals, and thus in populations different to potential target markets

4. There are many potential biases in real-world treatment comparisons and variations in data quality caused by acquisition in busy clinics

---

Fig. 1 The key differences between randomized controlled trials and real-world evidence studies [1, 3, 4]

---

and 3 MS studies, such as inflammatory lesions and relapse rates, are important surrogate markers for predicting the anticipated long-term prevention of disability, and these can only be validated via RWE methodologies.

Collection of high-quality real-world data (RWD) in MS

Regardless of disease area, collection of high-quality datasets in clinical practice requires agreement on a common minimum dataset and the use of special documentation software (or modifications to existing electronic medical records) to collect demographic and disease-specific information. These standards are important to offset the perception that the quality of RWD is inferior to that gathered in sophisticated, modern-day RCTs. Electronic medical record-based data collection is usually event based, with visits and other data being recorded as they occur, whereas RCTs use a schedule design where the timing and data collection at each visit are explicitly specified. In MS, event-based reporting currently predominates (e.g. the Swedish MS registry (SMSreg), the European Database for Multiple Sclerosis system, and

product launch, RWE can provide further insights into early clinical experience, safety, resource use, patient tolerability, and identification of untreated patients [3].
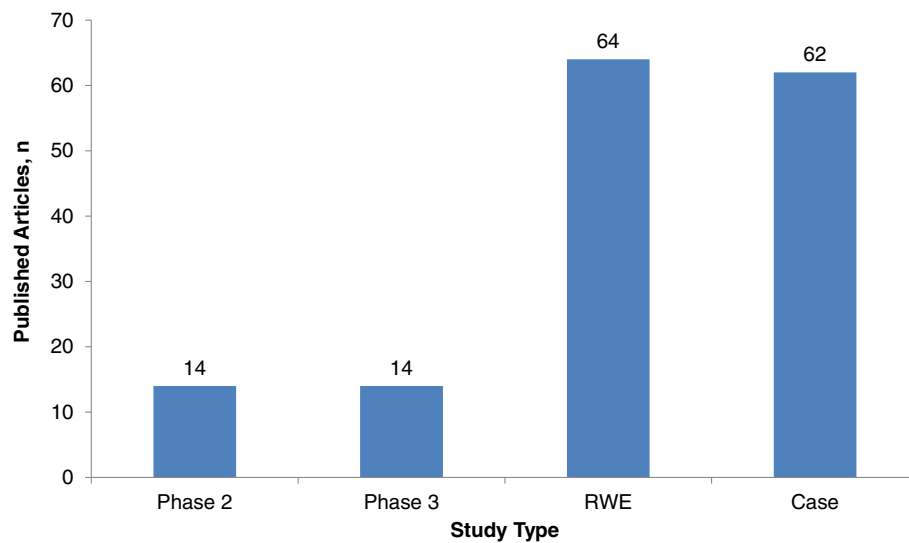
MS is a lifelong disease that can span more than 40 years. The short duration of RCTs provides limited information on MS disease course and long-term treatment effects. RCTs alone may be acceptable in acute neurological diseases like meningitis and stroke, where discrete endpoints, such as survival or post-acute fixed disability, can be swiftly measured. Conversely, for MS, the potential effects of a disease-modifying therapy (DMT) on disease progression can only be obtained by collecting information on patients treated in routine clinical practice over many years. The endpoints commonly used in phase 2

Ziemssen *et al. BMC Medicine* (2016) 14:81

Page 3 of 9

Fig. 2 Number of multiple sclerosis (MS) articles of different study types published in 2014. Numbers of each type of study were found by searching the PubMed database for articles published between January 1, 2014, and December 31, 2014, using the following search terms: for phase 2 clinical trials: ('multiple sclerosis' OR 'MS') AND ('Phase II clinical trials' OR 'Phase 2 clinical trials'); for phase 3 clinical trials: ('multiple sclerosis' OR 'MS') AND ('Phase III clinical trials' OR 'Phase 3 clinical trials'); For real-world evidence studies: ('multiple sclerosis' OR 'MS') AND ('real world' OR 'comparative effectiveness' OR 'registry'); for case studies: ('multiple sclerosis' OR 'MS') AND ('case study')

the iMED software used for data entry into the MS dataBase (MSBase)), although this is governed by specified minimum dataset descriptions. In contrast, the Multiple Sclerosis Documentation System 3D software used in multiple long-term follow-up studies imposes a defined visit schedule and is used predominantly in Germany, where it combines a 'trial-like' data documentation system with patient management [4]. Table 1 summarizes common sources of RWD [5] and gives examples of their use in RWE studies in MS [6–20].

Performing and identifying high-quality RWE studies in MS

In MS, RWD is increasingly used for comparison studies to examine therapy choice and sequencing decisions. Outcomes are similar to those in RCTs and include relapse and disability progression rates, adverse events and therapy discontinuation events. Identification and mitigation of biases and careful consideration of study power are key factors for designing appropriate RWE studies. Indeed, various biases exist and require careful consideration in selecting appropriate comparators, patient populations, data sources, outcomes and statistical analyses [21].

Selected analytical techniques (e.g. regression and stratification) can reduce the bias introduced by non-randomized study designs. Regression can improve the accuracy of an estimated treatment effect on a particular outcome measure by adjusting the association between treatment and outcome to account for other variables that could affect said outcome. In MS, these variables include

baseline Expanded Disability Status Scale (EDSS) score, prior relapse rate and disease duration. The type of regression model used largely depends on the outcome being measured [22]. Logistic regression is typically used for binary outcome measures (e.g. event occurred: yes/no), while proportional hazard estimates are used for continuous outcome measures (e.g. time to first relapse) [22]. Importantly, regression analysis uses all available data from the full patient cohort, enabling good statistical power, although it assumes that similar effects of an intervention occur across all subgroups and requires extrapolation when calculating estimates [22].

Stratification, whereby patient cohorts are divided into subgroups with similar variables, enables comparison of outcomes among patients with similar characteristics [22]. However, the potential for imbalance in other baseline covariates [22] requires the careful generation of statistically robust results, and sample sizes are necessarily smaller than for regression, reducing statistical power [22].

In many circumstances, propensity scoring has proven effective at reducing bias and is being used increasingly in longitudinal MS observational studies. In studies comparing the effect of two treatments, propensity scoring involves classification of the relationship between treatment assignment and baseline characteristics. Factors found to be different between treatment groups and associated with treatment choice are weighted to calculate the probability of any subject in the cohort being assigned a particular treatment – this is termed the 'propensity score'. Subjects are then matched by the

Ziemssen *et al. BMC Medicine* (2016) 14:81

Page 4 of 9

Table 1 Methodologies for collecting real-world data [4] and examples of their application to multiple sclerosis (MS) studies

Source	Explanation	Advantages	Limitations	Examples in MS
RCTs extensions	• Supplement trial data	• Extend RCT into real world	• Short study duration • Limited patient population • No information on rare events	• ENDORSE: EQ-5D and SF-36 in patients treated with DMF [10]
Registries	• Population-based collection of information	• Long-term data natural history and disease management • Regional comparisons	• Non-randomized design • Incomparable patient groupings • Discontinuous visit schedule • Varying practice patterns • Inter-regional extrapolation not always appropriate	• Lyons MS database: disability progression [11] • SMSreg: epidemiology and treatment outcomes [16, 19] • MSBase: treatment outcomes (69 countries) [13–15, 17]
Prospective observational studies	• Pre-defined outcome measures in clinical practice	• Robust dataset powered to answer specific questions	• Hawthorne effect (patients behave differently because they know they are being observed)	• PANGAEA, TOP: fingolimod and natalizumab clinical trial [9, 20]
PASSs	• Voluntary or imposed by regulatory authorities for approval	• Ongoing monitoring of the benefit–risk profile	• No obligation for regulatory submission of protocols and study reports for voluntary PASSs	• PANGAEA: German voluntary PASS [20]
Administrative data	• Data required for reimbursement	• Quick, low-cost analyses • Large patient populations provide information on rare events	• Privacy concerns limit access to data • Incomplete or inaccurate data • Costs and charges are not differentiated	• Pharmetrics Plus™ and Medco databases: relapses, treatment compliance, resource use and inpatient stays [6–8]
Health surveys	• Descriptive data	• Provide broadly generalizable data	• Not product-specific • Subjective • Relies on participant recollection	• NARCOMS survey: symptoms, comorbidities and health-related quality of life [18]
EMRs	• Real-time data collection	• Low cost • Detailed information over long periods	• High-end statistical analysis tools required	• EMRs: diagnosis, disease progression, symptoms and treatment [12]

DMF, Dimethyl fumarate; EMR, Electronic medical record; ENDORSE, BG00012 monotherapy safety and efficacy extension study in MS; EQ-5D, European Quality of Life-5 dimensions questionnaire; MSBase, Multiple Sclerosis dataBase; NARCOMS, North American Research Committee on Multiple Sclerosis; PANGAEA, Post-Authorization Noninterventional German sAfety of GilEnyA in RRMS patients; PASS, Post-authorization safety study; RCT, Randomized controlled trial; SF-36, 36-Item Short Form; SMSreg, Swedish MS registry; TOP, TYSABRI Observational Program

propensity score for comparison across treatment groups [23]. In effect, 'unmatchable' subjects, in whom a particular set of baseline characteristics leads to non-overlapping treatment assignment within a population, are removed from the outcome analysis and only 'matchable' subgroup outcomes are reported. Propensity scores can be used to aid stratification or regression through posterior adjustment of results [23]. When there are fewer than eight events per confounder, propensity scoring has been found to be a more robust method of eliminating bias than regression [24].

The introduction of bias through lack of randomization and blinding, as well as other methodological limitations of RWE studies, has raised questions about the validity of the evidence produced [25]. With this in mind, efforts have been made to promote good research practice and to advise researchers designing RWE studies to maximize the usefulness of the results obtained [21]. For cohort studies in general, checklists have been produced for both RCTs (CONSORT) [26] and for more general cohort studies (STROBE) [27]. Specifically for RWE studies, Dreyer et al. [28] recently compiled the Good Research for Comparative Effectiveness checklist to allow identification of RWE studies sufficiently high in quality for use in decision-making. Box 2 summarizes this checklist and the 10 'golden rules' for identifying high-quality RWE studies, all of which are relevant to study design in MS.

How has RWE helped in understanding the disease course and patient management in MS?

RWE generated from quality registries and other databases has greatly contributed to our understanding of MS disease course and risk factors. RWE studies have reported a decreased life expectancy for patients with MS compared with the general population [29], and have shown how factors such as increased age at disease onset and the primary progressive subtype of MS are associated with faster disability progression [30, 31]. They have also shown that, despite high variability in individual patients following conversion to secondary progressive MS and in patients with primary progressive MS, the mean or median progression rates between these disease subtypes are similar [31]. Other RWE studies have indicated a lower familial risk for developing MS than previously predicted [19], as well as an influence of race on outcomes [32].

RWE has also helped to guide patient management and treatment decisions in MS by answering questions related to treatment effects in clinical practice, which RCTs are unable to address. These effects are discussed in turn in the following sections.

Ziemssen *et al. BMC Medicine* (2016) 14:81

Page 5 of 9

Box 2. The ten golden rules for identifying high-quality real-world evidence (RWE) studies [28]

Rules for identifying high-quality RWE studies

1. Treatment details and primary outcomes should be adequately recorded
- Sufficient detail should be recorded so that information on treatment, including dose, regimen and mode of administration, can be determined
- Enough meaningful and robust information to allow primary outcomes of the study to be established

2. Primary outcomes should be measured objectively
- Objective primary outcome measures help to reduce bias between datasets
- Objective primary outcome measures for multiple sclerosis (MS) would include the number and volume of T2 lesions and disability progression as measured by the Expanded Disability Status Scale (EDSS)

3. Primary outcomes should be valid [28]
- The primary outcomes for the study should be appropriate for the research question to be answered (e.g. EDSS progression and relapse rate in MS, and claims data for resource use)

4. Primary outcomes should be identified and measured equally in treatment and comparison groups
- Data, tools and methods used for evaluating primary outcomes should be identified and measured in the same way in both the treatment and the comparison groups• Comparator arms should be clinically relevant [1]

5. Confounders of treatment effect should be recorded
- Any variables that could potentially impact the primary outcome of the study should be recorded (e.g. in MS, age of onset of disease, number of relapses in the year prior to start of the study, EDSS score at entry into study, previous treatment with disease-modifying therapies, and number of lesions at baseline)

6. The study population should be restricted to new users of the treatment being assessed
- Include only treatment-naïve patients to enable the frequency of early treatment effects to be determined and to limit bias associated with delayed treatment or loss of effectiveness
- If treatment naivety cannot be achieved, then washout should be considered

7. Data should be collected for the same time period for treatment and comparison groups
- Data from patients in treatment and comparator groups should be from within the same time period to ensure that the same standard of care is applied to both groups and that time-dependent effects are minimised
- Historical data may be used in some cases (e.g. when comparing with older treatments that are no longer widely used)

8. Confounders of treatment effect should be taken into account in the study design and analysis
- Studies should either restrict patient inclusion or employ analytical methods such as propensity scoring, stratification and regression to control for confounding factors
- Methods used for reducing bias should be properly reported; however, it is still important to be aware of the limitations of the statistical techniques used to reduce bias as they can lead to an inaccurate estimation of the effect of a treatment [1]

9. The study design should take into account 'immortal time bias'
- Study design should ensure that any events that occur before treatment exposure (in 'immortal time') are not included in the analysis, as this can exaggerate drug benefits

10. Key assumptions on which primary outcomes are based should be analyzed
- Assumptions should be analysed to verify their validity and reduce bias

---

Comparative effectiveness of DMTs in clinical practice

In the RWE setting, comparative effectiveness of DMTs can be determined from registry data. In MS, web-based registries, such as MSBase and SMSreg, can provide this type of information at global and regional levels [16, 33]. For instance, data from SMSreg have shown that fingolimod treatment initiation is associated with stable EDSS scores and improvement in relapse rates, disease severity, cognition, and quality of life after 12 months [16]. MSBase data analyses have also demonstrated the positive effect of DMT treatment on first confirmed disability progression [34] and have been used extensively for comparative effectiveness studies to show that, in patients who relapse on 'platform' injectable DMTs, switching to fingolimod or natalizumab, rather than between injectable DMTs, can lead to improvements in time to first relapse, relapse rate, and disability progression and regression events [17, 35]. Interestingly, after relapse on platform therapy,

Ziemssen *et al. BMC Medicine* (2016) 14:81

Page 6 of 9

a comparison of switch to fingolimod versus natalizumab indicated that switching to natalizumab from baseline therapies reduced relapse rates and increased sustained disability regression events more than switching to fingolimod, but that there was no difference in the rate of confirmed progression events between these treatments [14]. MSBase registry data also confirmed that relapse rates in patients switching to fingolimod from natalizumab were comparable to those switching from other therapies, and that an ideal treatment gap between these therapies was less than 8 weeks to reduce risk of early relapse [13] – a finding subsequently confirmed in an RCT [36]. MSBase data have also demonstrated variations between different baseline therapies, with patients treated with glatiramer acetate or subcutaneous interferon β-1α experiencing fewer relapses than those taking other baseline therapies [15]. Importantly, MSBase analyses (using propensity score matching) have generally replicated RCT results where known and align well with clinician experience, suggesting that biases are properly addressed using this dataset and statistical methodology.

Safety and tolerability of a DMT in clinical practice
Registries are also a good source of information on safety and tolerability of a product in a real-world setting. The Immunomodulation and Multiple Sclerosis Epidemiology studies used SMSreg data to show that, although fingolimod and natalizumab are both well tolerated, fingolimod tolerability is reduced compared with that of natalizumab, especially in patients switching from natalizumab [37]. The Post-Authorization Noninterventional German sAfety of GilEnyA in RRMS patients (PANGAEA) study is a prospective, observational, registry-based study designed to collect data on effectiveness and adverse events in fingolimod-treated patients in standard clinical practice [20]. Prospective observational studies performed without registry data also provide an insight into the safety and tolerability of MS treatments in the real world, albeit on a smaller scale. The Safety, Tolerability and Adherence with Rebif© study and Tysabri Observational Program interim analysis, for example, demonstrated that the safety profiles of subcutaneous interferon β-1α and natalizumab in relapsing-remitting MS patients in clinical practice are comparable with those in RCTs [9, 38].

Impact of a DMT on resource use and medical costs
Claims databases provide valuable data on the impact of DMTs on resource use and medical costs, such as use of other medications and hospital stays. These pharmacoeconomic considerations are important influencing factors in formulary decisions. IMS PharMetrics Plus™ is a medical and pharmacy claims database containing the records of over 100,000 patients with MS across the USA [7, 8]. This dataset has been used to distinguish the

effects of different DMTs on outcome measures such as adherence, persistence, inpatient stays and corticosteroid use for relapses [7, 8].

Impact of a DMT on pharmacoeconomic outcomes
Pharmacoeconomic studies provide an insight into the cost-effectiveness of a DMT, including the direct effect of medication costs and indirect costs from health-related work absences or caregiver time. The PANGAEA and ProspEctive phArmacoeconomic cohoRt evaLuation (PEARL) registry studies of fingolimod- and injectable DMT-treated patients collect data on sick leave, hospitalization and physician consultations, and show that fingolimod treatment results in better pharmacoeconomic outcomes compared with baseline therapies [39]. Similar results have been observed with natalizumab treatment in patients in the USA through analysis of administrative claims databases [40].

Impact of a DMT on patient-reported outcomes (PROs)
Patient-generated data can be used to assess the perceived clinical benefits of a therapy and is useful for examining other aspects of patient experiences with DMTs. Some RCTs, such as the Evaluate Patient Outcomes study [41], which assessed treatment satisfaction in patients treated with fingolimod or injectable DMTs, collect patient-generated data. Some observational RWE studies also assess PROs. PANGAEA and PEARL were designed to capture patient experience on effectiveness, tolerability [20] and treatment satisfaction, as well as on ease and convenience of taking their DMT as instructed [20, 42]. Additional PROs on quality of life, physical disability, cognition and fatigue can be obtained through health surveys [43].

Evolving RWD collection in MS
Many countries across the world now collect MS RWD in registries and other databases [30, 44] (Table 2), and sharing of this information is critical to increase statistical power and to enable inter-regional knowledge transfer. Collaborations between MS registries, such as the European Register for Multiple Sclerosis, which includes data from 13 registries in Europe [45], further highlight the importance of information-sharing and learning on an international level by providing expanded datasets that allow cross-border analysis, interpretation and dissemination of results. Another collaborative example is the BigMS collaboration between MSBase and registries in France, Italy, Sweden and Denmark, which contains quality longitudinal datasets for over 130,000 patients with MS, with the first joint analyses planned for 2016. However, while the sharing of registry data sounds like an obvious positive step, in some cases this is not possible because it may breach the original patient consent. In these cases, efforts should be made to either re-contact the patients or to strip out all patient

Ziemssen *et al. BMC Medicine* (2016) 14:81

Page 7 of 9

Table 2 Key multiple sclerosis (MS) registry studies in Europe, North America and globally

Registry	Patients, n
Europe	
Croatia (SDMSH)	2477
Denmark (DMSC)	12,500
France (EDMUS)	~40,000
Germany (DMSG)	~30,000
Germany (MSDS3D Users)	>5000
Greece (GMSS)	3500
Italy (iMed/web)	~20,000
Norway (Nasjonal kompetansetjeneste for multippel sklerose – MS)	5100
Russia	21,500
Serbia	3500
Spain (Catalonia; EpidEMcat)	>5000
Sweden (SMSreg)	12,900
Switzerland (SMSR)	270
United Kingdom (MS Register)	8300
North America	
Canada (London Ontario)	1099
Canada (British Columbia)	2837
USA (New York; NYSMSC)	9600[a]
USA (NARCOMS)	19,297
Global	
MSBase (31 countries)	39,030

Patient numbers are derived from the original publications [30, 44] and may have changed
[a]Numbers updated by NYSMSC, 6 May 2016
DMSC, Danish Multiple Sclerosis Center; DMSG, National Multiple Sclerosis Society of Germany; EDMUS, European Database for Multiple Sclerosis; EpidEMcat, Catalonia MS Registry; GMSS, Greek Multiple Sclerosis Society; NARCOMS, North American Research Committee on Multiple Sclerosis; NYSMSC, New York State Multiple Sclerosis Consortium; MSBase, Multiple Sclerosis database; SDMSH, Croatian Multiple Sclerosis Society; SMSR, Scottish Multiple Sclerosis Register; SMSreg, Swedish MS registry

identifiers before the data are shared. Certainly, with the modem age of big data and the need to combine datasets to improve statistical power, consent to share data ought to be included in the consent forms for any new programmes.

MS RWD can be collected within registries that contain information on other disorders that can be used to inform neurologists and improve patient management across a range of neurologic diseases. The Swedish Neuro Registries includes MS RWD from SMSreg as well as other disease-specific data, including Parkinson's disease, epilepsy and myasthenia gravis. However, disease specificity and usability of registry data is fundamentally assured by the registry participant's agreement to collect a uniformly defined 'minimum dataset' at a relatively set frequency. This is a major limitation of MS registries, and if data are to become more reliable and powerful, then standards for data input are needed and these need to be audited with validated quality control checks.

Although registries currently collect information on a variety of different clinical, pharmacoeconomic and safety outcome measures, the quality of magnetic resonance imaging (MRI) data in MS registries is currently limited. The descriptive, semi-quantitative MRI T2 lesion measures that are often recorded (e.g. Barkhof–Tintoré criteria) capture the severity of cerebral and spinal cord MRI abnormalities in MS relatively poorly, although they do provide a broad characterisation of lesion load and location.

The argument for inclusion of better, standardised MRI data in registries, and indeed their acquisition in clinical practice, is strong. MRI provides evidence of disease activity in MS [46] and new T2 hyperintense and T1 hypointense lesion development, particularly whilst on treatment, correlates with long-term disability progression [47, 48]. The increased rate of brain volume (BV) loss in patients with MS compared with healthy controls [49], indicating myelin, axonal and neuronal loss [50], is associated with cognitive impairment and disability in MS [51–53]. As BV loss is differentially affected by different MS therapies, standardized measures of BV change in clinical practice captured in real-world registries could provide information on the relative efficacy of MS therapies [54], with such information warranting inclusion in prognostic models. Inclusion of quantitative baseline MRI parameters would also improve patient matching to compare different treatment switch decisions in long-term RWE studies. We believe that a technological solution is imminent. Indeed, third party providers performing online automated MRI analytics (e.g. IcoMetrix, NeuroQuant) are emerging and have achieved medical device registration for use in clinical practice. The MRI manufacturers themselves are also developing automated lesion and BV algorithms that will likely be incorporated in standard MRI analytical packages within the next 2 years.

Importantly, RWE collection systems are becoming increasingly flexible, with the ability to exchange data with different sources (e.g. physician-, MRI- and patient-generated data) [55]. Additionally, key providers of RWD collection systems understand the need for creating added value for contributing clinicians, with features like graphical outcomes recording, own data exportation and data benchmarking analysis being offered through the platforms.

Conclusions

RWE provides a valuable contribution to the evidence base for the use of MS therapies by supplementing RCT data and providing information on long-term effectiveness and tolerability of treatments in a real-world setting across generalizable populations. It provides longitudinal outcome information that can be directly used for pharmacoeconomic evaluation and examination of treatment patterns and sequencing outcomes. Robust study designs,

Ziemssen *et al. BMC Medicine* (2016) 14:81

Page 8 of 9

including appropriate data collection and analytical methods as well as a unified minimum dataset, are critical for generating evidence that can be reasonably used to influence treatment decisions and guidelines, and satisfy the growing need of stakeholders to monitor DMT performance post-approval.

RWE generation can be further expanded in MS, as some important outcome measures are not routinely collected in real-world databases. By collecting information on additional outcome measures, such as MRI and cognition, their importance in guiding treatment could be examined. By continually expanding the data pool through collaborations, the validity and utility of RWE to physicians and regulatory bodies will be improved, fostering greater and better physician/patient engagement.

Abbreviations
BV: brain volume; DMT: disease-modifying therapy; EDSS: Expanded Disability Status Scale; MRI: magnetic resonance imaging; MS: multiple sclerosis; MSBase: multiple sclerosis database; PANGAEA: Post-Authorization Noninterventional German sAfety of GilEnyA in RRMS patients; PEARL: prospEctive phArmacoeconomic cohoRt evaluation; PRO: patient-reported outcome; RCT: randomized controlled trial ; RWD: real-world data; RWE: real-world evidence; SMSreg: Swedish MS registry.

Authors' contributions
TZ participated in study conception and design and drafting of the manuscript. All authors performed the acquisition, analysis and interpretation of the data, critically revised the manuscript for important intellectual content, and approved the final version of the manuscript.

Competing interests
TZ has received compensation for consulting services from Almirall, Biogen Idec, Bayer, Genzyme, GlaxoSmithKline, MSD, Merck Serono, Novartis, Sanofi, Teva, and Synthon, and has received research support from Bayer, Biogen Idec, the Hertie Foundation, the Roland Ernst Foundation, the German Diabetes Foundation, Merck Serono, Novartis, Teva, and Sanofi Aventis. Further, he is a lead investigator in the PANGAEA and PEARL study.
JH has received honoraria for serving on advisory boards for Biogen Idec, Genzyme, and Novartis, and has received speaker's fees from Biogen Idec, Merck Serono, Bayer-Schering, Teva, Novartis, and Sanofi-Genzyme. He has served as Principal Investigator for projects sponsored by, or received unrestricted research support from, Biogen Idec, Merck Serono, Teva, Novartis, and Bayer-Schering. His MS research is funded by the Swedish Research Council and the Swedish Brain Foundation.
HB has served on scientific advisory boards for Biogen Idec, Novartis, Genzyme and Merck, and has received conference travel support from Novartis, Biogen, and Genzyme. He serves on steering committees for trials conducted by Biogen and Novartis, and his institution has received research support from Merck, Novartis, CSL Biopharma, and Biogen. He has received research support from MS Research Australia, Charityworks for MS, the National Health and Medical Research Council of Australia, and the Royal Melbourne Hospital Neuroscience Foundation.

Author details
[1]Center of Clinical Neuroscience, Department of Neurology, MS Center Dresden, Center of Clinical Neuroscience, University Hospital Carl Gustav Carus, Dresden University of Technology, Fetscherstr. 74, 01307 Dresden, Germany. [2]Department of Clinical Neuroscience and Center for Molecular Medicine, Karolinska Institute, Stockholm, Sweden. [3]Department of Neurology, Royal Melbourne Hospital, Victoria, Australia.

References
1. Nallamothu BK, Hayward RA, Bates ER. Beyond the randomized clinical trial: the role of effectiveness studies in evaluating cardiovascular therapies. Circulation. 2008;118:1294–303.
2. Milnes F, Bergvall N, Olson M, Capkun-Niggli G, Bonzani I, Makin C. Use of real-world data to assess outcomes in patients with multiple sclerosis in a real-life clinical setting. 10th Annual Meeting of Health Technology Assessment international (HTAi). Seoul: Poster P1082; 2013.
3. Association of the British Pharmaceutical Industry. Demonstrating value with real world data: a practical guide. 2011. http://www.abpi.org.uk/our-work/library/guidelines/Pages/real-world-data.aspx. Accessed 25 February 2016.
4. Garrison Jr LP, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: the ISPOR Real-World Data Task Force report. Value Health. 2007;10:326–35.
5. Ziemssen T, Kempcke R, Eulitz M, Grossmann L, Suhrbier A, Thomas K, Schultheiss T. Multiple sclerosis documentation system (MSDS): moving from documentation to management of MS patients. J Neural Transm (Vienna). 2013;120 Suppl 1:S61–6.
6. Agashivala N, Wu N, Abouzaid S, Wu Y, Kim E, Boulanger L, Brandes DW. Compliance to fingolimod and other disease modifying treatments in multiple sclerosis patients, a retrospective cohort study. BMC Neurol. 2013;13:138.
7. Bergvall N, Lahoz R, Reynolds T, Korn JR. Healthcare resource use and relapses with fingolimod versus natalizumab for treating multiple sclerosis: a retrospective US claims database analysis. Curr Med Res Opin. 2014;30:1461–71.
8. Bergvall N, Petrilla AA, Karkare SU, Lahoz R, Agashivala N, Pradhan A, Capkun G, Makin C, McGuiness CB, Korn JR. Persistence with and adherence to fingolimod compared with other disease-modifying therapies for the treatment of multiple sclerosis: a retrospective US claims database analysis. J Med Econ. 2014;17:696–707.
9. Butzkueven H, Kappos L, Pellegrini F, Trojano M, Wiendl H, Patel RN, Zhang A, Hotermans C, Belachew S; TYSABRI Observational Program (TOP) Investigators. Efficacy and safety of natalizumab in multiple sclerosis: interim observational programme results. J Neurol Neurosurg Psychiatry. 2014;85:1190–7.
10. A dose-blind, multicenter, extension study to determine the long-term safety and efficacy of two doses of BG00012 monotherapy in subjects with relapsing-remitting multiple sclerosis. ClinicalTrials.gov Identifier: NCT00835770. 2015. https://clinicaltrials.gov/show/NCT00835770. Accessed 25 February 2016.
11. Confavreux C, Vukusic S, Moreau T, Adeleine P. Relapses and progression of disability in multiple sclerosis. N Engl J Med. 2000;343:1430–8.
12. Davis MF, Sriram S, Bush WS, Denny JC, Haines JL. Automated extraction of clinical traits of multiple sclerosis in electronic medical records. J Am Med Inform Assoc. 2013;20:e334–40.
13. Jokubaitis VG, Li V, Kalincik T, Izquierdo G, Hodgkinson S, Alroughani R, Lechner-Scott J, Lugaresi A, Duquette P, Girard M, et al. Fingolimod after natalizumab and the risk of short-term relapse. Neurology. 2014;82:1204–11.
14. Kalincik T, Horakova D, Spelman T, Jokubaitis V, Trojano M, Lugaresi A, Izquierdo G, Rozsa C, Grammond P, Alroughani R, et al. Switch to natalizumab versus fingolimod in active relapsing-remitting multiple sclerosis. Ann Neurol. 2015;77:425–35.
15. Kalincik T, Jokubaitis V, Izquierdo G, Duquette P, Girard M, Grammond P, Lugaresi A, Oreja-Guevara C, Bergamaschi R, Hupperts R, et al. Comparative effectiveness of glatiramer acetate and interferon beta formulations in relapsing-remitting multiple sclerosis. Mult Scler. 2015;21:1159–71.
16. Nordin N, Hillert J, Piehl F, Svenningsson A, Lycke J, Fagius J, Walentin F, Martin C, Nilsson P, Feltelius N, et al. The "Immunomodulation and Multiple Sclerosis Epidemiology" (IMSE II) study; A Swedish nationwide pharmaco-epidemiological and genetic study focused on long-term safety and efficacy of fingolimod (Gilenya®), 29th Congress of the European Committee for Research and Treatment in Multiple Sclerosis (ECTRIMS). Copenhagen: Poster; 2013. p. 975.

Ziemssen *et al. BMC Medicine* (2016) 14:81

Page 9 of 9

17. He A, Spelman T, Jokubaitis V, Havrdova E, Horakova D, Trojano M, Lugaresi A, Izquierdo G, Grammond P, Duquette P, et al. Comparison of switch to fingolimod or interferon beta/glatiramer acetate in active multiple sclerosis. JAMA Neurol. 2015;72:405–13.
18. North American Research Committee on Multiple Sclerosis.. Publications validating NARCOMS data and instruments. 2016. http://narcoms.org/publications. Accessed 26 February 2016.
19. Westerlind H, Ramanujam R, Uvehag D, Kuja-Halkola R, Boman M, Bottai M, Lichtenstein P, Hillert J. Modest familial risks for multiple sclerosis: a registry-based study of the population of Sweden. Brain. 2014;137:770–78.
20. Ziemssen T, Kern R, Cornelissen C. The PANGAEA study design – a prospective, multicenter, non-interventional, long-term study on fingolimod for the treatment of multiple sclerosis in daily practice. BMC Neurol. 2015;15:93.
21. Willke RJ, Mullins CD. Ten commandments" for conducting comparative effectiveness research using "real-world data. J Managed Care Pharm. 2011;17:S10–5.
22. Normand SL, Sykora K, Li P, Mamdani M, Rochon PA, Anderson GM. Readers guide to critical appraisal of cohort studies: 3. Analytical strategies to reduce confounding. BMJ. 2005;330:1021–3.
23. Austin PC. An Introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res. 2011;46:399–424.
24. Cepeda MS, Boston R, Farrar JT, Strom BL. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. Am J Epidemiol. 2003;158:280–7.
25. Möller HJ. Effectiveness studies: advantages and disadvantages. Dialogues Clin Neurosci. 2011;13:199–207.
26. CONSORT. The CONSORT Statement. 2010. http://www.consort-statement.org/. Accessed 12 April 2016.
27. STROBE. Checklist for cohort studies. 2007. http://www.strobe-statement.org/index.php?id=available-checklists. Accessed 12 April 2016.
28. Dreyer NA, Velentgas P, Westrich K, Dubois R. The GRACE checklist for rating the quality of observational studies of comparative effectiveness: a tale of hope and caution. J Manag Care Spec Pharm. 2014;20:301–8.
29. Hurwitz BJ. Analysis of current multiple sclerosis registries. Neurology. 2011;76:S7–13.
30. Hurwitz BJ. Registry studies of long-term multiple sclerosis outcomes: description of key registries. Neurology. 2011;76:S3–6.
31. Kremenchutzky M, Rice GP, Baskerville J, Wingerchuk DM, Ebers GC. The natural history of multiple sclerosis: a geographically based study 9: observations on the progressive phase of the disease. Brain. 2006;129:584–94.
32. Weinstock-Guttman B, Jacobs LD, Brownscheidle CM, Baier M, Rea DF, Apatoff BR, Blitz KM, Coyle PK, Frontera AT, Goodman AD, et al. Multiple sclerosis characteristics in African American patients in the New York State Multiple Sclerosis Consortium. Mult Scler. 2003;9:293–8.
33. Butzkueven H, Chapman J, Cristiano E, Grand'Maison F, Hoffmann M, Izquierdo G, Jolley D, Kappos L, Leist T, Pohlau D, et al. MSBase: an international, online registry and platform for collaborative outcomes research in multiple sclerosis. Mult Scler. 2006;12:769–74.
34. Jokubaitis VG, Spelman T, Kalincik T, Izquierdo G, Grand'Maison F, Duquette P, Girard M, Lugaresi A, Grammond P, Hupperts R, et al. Predictors of disability worsening in clinically isolated syndrome. Ann Clin Transl Neurol. 2015;2:479–91.
35. Spelman T, Kalincik T, Zhang A, Pellegrini F, Wiendl H, Kappos L, Tsvetkova L, Belachew S, Hyde R, Verheul F, et al. Comparative efficacy of switching to natalizumab in active multiple sclerosis. Ann Clin Transl Neurol. 2015;2:373–87.
36. Kappos L, Radue EW, Comi G, Montalban X, Butzkueven H, Wiendl H, Giovannoni G, Hartung HP, Derfuss T, Naegelin Y, et al. Switching from natalizumab to fingolimod: a randomized, placebo-controlled study in RRMS. Neurology. 2015;85:29–39.
37. Frisell T, Forsberg L, Nordin N, Kiesel C, Alfredsson L, Askling J, Hillert J, Olsson T, Piehl F. Comparative analysis of first-year fingolimod and natalizumab drug discontinuation among Swedish patients with multiple sclerosis. Mult Scler. 2016;22:85–93.
38. Hupperts R, Ghazi-Visser L, Martins Silva A, Arvanitis M, Kuusisto H, Marhardt K, Vlaikidis N, Group SS. The STAR Study: a real-world, international, observational study of the safety and tolerability of, and adherence to, serum-free subcutaneous interferon beta-1a in patients with relapsing multiple sclerosis. Clin Ther. 2014;36:1946–57.
39. Ziemssen T, Moser S, Schwartz HJ, Fuchs A, van Lockven T, Cornelissen C. Comparison of efficacy and treatment satisfaction of patients on baseline

therapy and fingolimod: results of the interim analysis of two German registry studies PANGAEA and PEARL. American Academy of Neurology (AAN) Annual Meeting, April 26–May 3, 2014, Philadelphia PA; Poster P7.232.
40. Bonafede MM, Johnson BH, Watson C. Health care-resource utilization before and after natalizumab initiation in multiple sclerosis patients in the US. Clinicoecon Outcomes Res. 2013;6:11–20.
41. Fox E, Edwards K, Burch G, Wynn DR, LaGanke C, Crayton H, Hunter SF, Huffman C, Kim E, Pestreich L, et al. Outcomes of switching directly to oral fingolimod from injectable therapies: Results of the randomized, open-label, multicenter, Evaluate Patient OutComes (EPOC) study in relapsing multiple sclerosis. Mult Scler Relat Disord. 2014;3:607–19.
42. Vormfelde SV, Ortler S, Ziemssen T. Multiple Sclerosis Therapy With Disease-Modifying Treatments in Germany: The PEARL (ProspEctive phArmacoeconomic cohoRt evaluation) Noninterventional Study Protocol. JMIR Res Protocol. 2015;5(1), e23.
43. Stephenson JJ, Kern DM, Agarwal SS, Zeidman R, Rajagopalan K, Kamat SA, Foley J. Impact of natalizumab on patient-reported outcomes in multiple sclerosis: a longitudinal study. Health Qual Life Outcomes. 2012;10:155.
44. Flachenecker P, Buckow K, Pugliatti M, Kes VB, Battaglia MA, Boyko A, Confavreux C, Ellenberger D, Eskic D, Ford D, et al. Multiple sclerosis registries in Europe - results of a systematic survey. Mult Scler. 2014;20:1523–32.
45. A collaborative initiative to improve MS research and policy across Europe. Better outcomes with better data. 2014. http://www.emsp.org/wp-content/uploads/2015/06/140930-EUReMS-Report.pdf. Accessed 26 February 2016.
46. Barkhof F, Scheltens P, Frequin ST, Nauta JJ, Tas MW, Valk J, Hommes OR. Relapsing-remitting multiple sclerosis: sequential enhanced MR imaging vs clinical findings in determining disease activity. AJR Am J Roentgenol. 1992;159:1041–7.
47. Giorgio A, Stromillo ML, Bartolozzi ML, Rossi F, Battaglini M, De Leucio A, Guidi L, Maritato P, Portaccio E, Sormani MP, et al. Relevance of hypointense brain MRI lesions for long-term worsening of clinical disability in relapsing multiple sclerosis. Mult Scler. 2014;20:214–9.
48. Sormani MP, Rio J, Tintore M, Signori A, Li D, Cornelisse P, Stubinski B, Stromillo M, Montalban X, De Stefano N. Scoring treatment response in patients with relapsing multiple sclerosis. Mult Scler. 2013;19:605–12.
49. Barkhof F, Calabresi PA, Miller DH, Reingold SC. Imaging outcomes for neuroprotection and repair in multiple sclerosis trials. Nat Rev Neurol. 2009;5:256–66.
50. Simon JH. Brain atrophy in multiple sclerosis: what we know and would like to know. Mult Scler. 2006;12:679–87.
51. Calabrese M, Agosta F, Rinaldi F, Mattisi I, Grossi P, Favaretto A, Atzori M, Bernardi V, Barachino L, Rinaldi L, et al. Cortical lesions and atrophy associated with cognitive impairment in relapsing-remitting multiple sclerosis. Arch Neurol. 2009;66:1144–50.
52. Sormani MP, Arnold DL, De Stefano N. Treatment effect on brain atrophy correlates with treatment effect on disability in multiple sclerosis. Ann Neurol. 2014;75:43–9.
53. Prinster A, Quarantelli M, Lanzillo R, Orefice G, Vacca G, Carotenuto B, Alfano B, Brunetti A, Morra VB, Salvatore M. A voxel-based morphometry study of disease severity correlates in relapsing– remitting multiple sclerosis. Mult Scler. 2010;16:45–54.
54. Zivadinov R, Bakshi R. Central nervous system atrophy and clinical status in multiple sclerosis. J Neuroimaging. 2004;14:27S–35S.
55. Kern R, Haase R, Eisele JC, Thomas K, Ziemssen T. Designing an Electronic Patient Management System for Multiple Sclerosis: Building a Next Generation Multiple Sclerosis Documentation System. Interact J Med Res. 2016;5(1), e2.