

Time for Change – Evolution of Real-world Evidence Outcome Measures in Multiple Sclerosis Exemplified by Fingolimod

Report of a Satellite Symposium Held at the Joint Congress of European Neurology, Istanbul, Turkey, 1 June 2014

Expert review by: Mefkure Eraksoy,¹ Helmut Butzkueven,² Tjalf Ziemssen³ and Robert Zivadinov^{4,5}

1. Istanbul Faculty of Medicine, Istanbul University, Department of Neurology, Istanbul, Turkey; 2. Director, MS Service, Royal Melbourne and Box Hill Hospitals, Melbourne, Australia; Chair, MSBase Foundation, Department of Neurology, Royal Melbourne Hospital, Victoria, Australia; 3. Multiple Sclerosis Centre, Centre of Clinical Neuroscience, University Hospital Carl Gustav Carus, Dresden, Germany; 4. Department of Neurology, Buffalo Neuroimaging Analysis Center, School of Medicine and Biomedical Sciences, University at Buffalo, State University of New York, NY, US; 5. Professor of Neurology; Director, Buffalo Neuroimaging Analysis Center; Director, MRI Clinical Translational Research Center, School of Medicine and Biomedical Sciences, University at Buffalo, State University of New York, NY, US

Abstract

Real-world evidence provides important information concerning the long-term effectiveness and safety of disease-modifying treatments (DMTs) for multiple sclerosis (MS) in clinical practice in a large number of patients. These data enhance and extend the results from randomised clinical trials and include information that cannot easily be obtained in trials such as treatment efficacy in non-trial populations (e.g. those with co-morbidities, older patients and children) and long-term safety analyses. Such data are compatible with specialist clinical practice and, when accumulated in multicentre databases using an agreed minimum dataset, these data can provide invaluable information on long-term disease progression and treatment. Real-world evidence in terms of fingolimod is reviewed and clearly demonstrates its long-term effectiveness and safety. Currently, imaging data are not systematically collected but the development of improved standard magnetic resonance imaging (MRI) protocols and automated procedures now makes this a realistic option in the near future. Automated (or largely automated) MRI analytics are likely to greatly enhance and further strengthen real-world evidence for outcomes of different treatment algorithms incorporating more sophisticated MRI measures.

Keywords

Real-world data, real-world evidence, databases, registries, effectiveness, fingolimod, multiple sclerosis

Disclosure: Mefkure Eraksoy is a consultant for and has received travel support from Bayer Healthcare-Turkey, Biogen-Idec, Gen Pharmaceuticals-Turkey, Merck, Novartis, Serono-Turkey and TEVA Pharmaceuticals-Turkey, and is a speaker for Genzyme. Helmut Butzkueven has served on scientific advisory boards and has acted as a consultant for Biogen Idec and Novartis; has been a speaker for Biogen Idec Australia, Merck Serono Australia and Novartis Australia; and received travel support from Biogen Idec Australia and Merck Serono Australia and research support from CASS Foundation Australia, Merck Serono Australia, the Royal Melbourne Hospital Friends of Neurosciences Foundation and the University of Melbourne. Tjalf Ziemssen has served on scientific advisory boards and acted as a consultant for Bayer, Biogen Idec, Genzyme, Novartis and TEVA and has received speaker honoraria from Biogen Idec, Merck Serono and Novartis. Robert Zivadinov has been a speaker and consultant for Biogen Idec, Claret Medical, Genzyme-Sanofi, Novartis and Serono; has received research support from Biogen Idec, Claret Medical, EMD Serono, Genzyme-Sanofi, Greatbatch, Novartis, InteKrin, TEVA Pharmaceuticals, the US Department of Defense, the US National Institutes of Health, the US National Science Foundation and the US National Multiple Sclerosis Society.

Acknowledgement: Medical writing support was provided by Ray Ashton at Touch Medical Media, London, UK.

Received: 16 July 2014 **Accepted:** 24 November 2014 **Citation:** *European Neurological Review*, 2014;9(2):136–42. DOI:10.17925/ENR.2014.09.02.136

Correspondence: Mefkure Eraksoy, Istanbul Faculty of Medicine, Istanbul University, Department of Neurology, Istanbul, Turkey. E: mefkure.eraksoy@superonline.com

Support: The publication of this article was supported by Novartis Pharmaceuticals. Novartis Pharmaceuticals' involvement in the development of this article was strictly limited to ensuring the accuracy of the data.

Relevance of Systematic Real-world Data Collection for Physicians and Patient

Physicians routinely collect data from multiple sclerosis (MS) patients as part of clinical practice. Much of these data are now in electronic formats. Minimum datasets for monitoring clinical outcomes are similar in many centres, and consensus is generally easily reached. When aggregated, these data potentially produce a powerful database. When paper records were ubiquitous, clinical research was largely separate from clinical practice with clinical research organisations and pharma companies populating databases via questionnaires. The development of electronic health records

is a major improvement, but an agreed minimum dataset and data-sharing arrangements are also required to monitor a long-term disease such as MS. The information in electronic medical records is 'big data' and typically describe overall global trends, but are not necessarily suitable to analyse individual patient outcomes. In common with a clinical trial, a disease registry (or long-term database) uses an agreed minimum dataset. Examples in MS include the European Database for Multiple Sclerosis (EDMUS) project,¹ the Swedish National Registry,² the Danish National Registry³ and the global MS Registry (MSBase).⁴ Once the language (i.e. the minimum dataset) has been agreed, typically including demographics, disease classification, relapse

dates, Expanded Disability Status Scale (EDSS) scores, disease-modifying drug (DMD) start and stop dates, codified magnetic resonance imaging (MRI) and other diagnostic testing, it can be implemented as a physician-initiated registry. Ideally, the registry data collection effort is highly integrated with clinical practice, with little extra time taken. A common misconception among physicians is lack of time, but the large existing registries with over 150,000 MS records in aggregate illustrate that it is possible with proper motivation and, ideally, some dedicated resources. For example, the Danish registry was started in 1948 and now includes over 25,000 records.³ It demonstrates the motivation of the Danish neurologists with minimal additional resources to produce this valuable dataset.

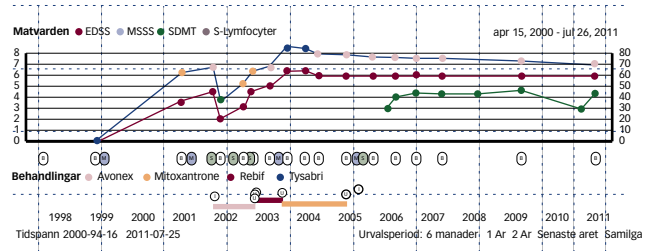
Most of the successful approaches are hybrids – comprising both a registry function and an e-health record function. For example, 80 % of Swedish MS patients are monitored online using such a system that provides data to the clinician on treatment, relapses and EDSS scores, together with comparative information on the severity of disease compared with other patients (see Figure 1).² Registries provide long-term information displayed graphically making it easy to relate the patients' current situation to the past, thus helping their management.

The Serono Symposia International Foundation originated the MSBase registry in 2000 to collect codified data with ethical approval and patient consent. The project was relatively unsuccessful due to perceived direct pharmaceutical company influence, and thus became independent in 2004 registered as a not-for-profit organisation in Australia. It has now registered over 31,500 patients, with recruitment significantly accelerating in the 2013–2014 period. In total 28 countries contribute to the registry, involving more than 150 participating neurologists, and the patient datasets have generated over 165,000 patient-years of follow-up. Median patient follow-up is 5 years and EDSS scores are submitted regularly (median 4–5 months), a data density equivalent to some clinical trials. One advantage of the Swedish Registry and MSBase is their built-in capacity to benchmark patient outcomes, and in the case of MSBase the 'MS severity calculator' is open to non-members of the system online. These MSBase severity calculators summarise EDSS score rank at various disease durations by displaying 25th, 50th and 75th percentiles using the entire MSBase dataset as a 'live' reference population. EDSS scores of an individual patient can be entered and the severity calculator provides percentile of the EDSS for any given annualised disease duration (see Figure 2). In the future, more sophisticated benchmarking efforts could be utilised to analyse potential clinical therapeutic decision outcomes, especially if MRI metrics can be incorporated.

Randomised clinical trials in MS provide information on the comparative efficacy (relapses, EDSS, MRI lesions, brain volume loss) of a disease-modifying treatment (DMT) in a controlled setting usually lasting up to 2 years.⁵ By contrast, real-world data assess the effectiveness of a DMT in a real-world setting over a longer period. There are many kinds of real-world data including MS registries giving comparative effectiveness (relapses, EDSS, treatment discontinuation) of a DMT in clinical practice, claims databases illustrating the impact of a DMT on resource utilisation and medical costs, population-based registries giving the impact of a DMT on clinical, safety, resource use and socioeconomic outcomes and observational studies that usually monitor a single DMT over a long period reporting its impact on multiple patient and clinical-reported outcomes.¹⁻³ In MS a number of important questions can potentially be addressed with real-world data:

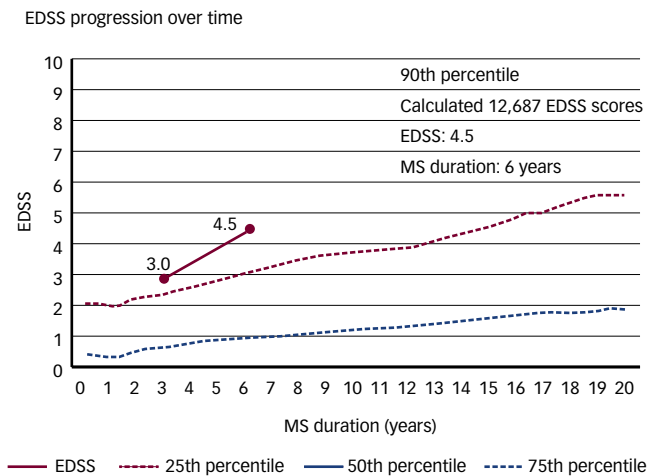
- How does DMT compare head-to-head if no randomised clinical trial has been reported?

Figure 1: Swedish Online Registry of MS Patients Combining Registry and E-health Record Functions



EDSS = expanded disability status scale; MS = multiple sclerosis; SDMT = symbol digit modalities test. Source: www.neuroreg.se

Figure 2: MSBase – Severity Calculator Based on MS Duration and Expanded Disability Status Scale Scores

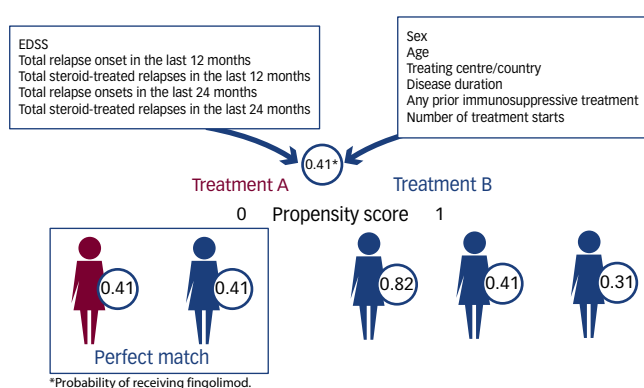


EDSS = Expanded Disability Status Scale; MS = Multiple Sclerosis. Source: <https://www.msbase.org/>

- Do DMTs prevent long-term disability accumulation?
- What is the impact of relapses on long-term disability progression?
- What is the long-term disability impact of pregnancy? Do relapses in pregnancy influence long-term disability?
- Do DMTs work in progressive forms of MS?
- Can treatment be individualised to maximise benefit:risk ratios for people with MS?

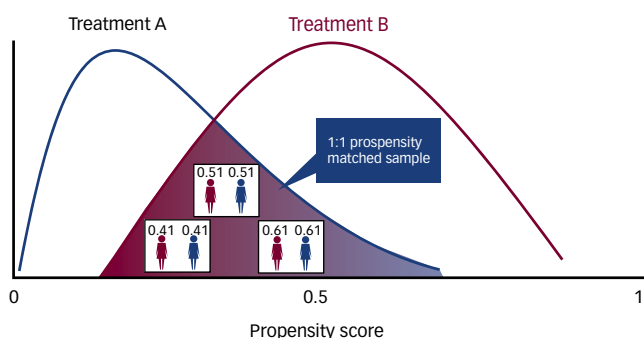
In terms of treatment comparisons, a potential problem in the real world is the lack of randomisation so that baseline variables between different treatment groups may not be well matched. This problem can be addressed by various techniques, and one popular method is propensity score matching. Individual patients with different treatment decisions (for instance, switch to fingolimod or natalizumab after relapse on injectable DMD compared with switch between injectable DMD classes) can be selected as pairs on the basis of closely 'matched' baseline characteristics, using both patients demographic and prior disease activity/severity.⁶ The statistical method merges all significant predictors of differential treatment assignment into a single score, known as a propensity score, representing the conditional probability that a particular patient receives one of the two treatment options. For example, a patient in group A having a conditional propensity score of 0.41 of receiving treatment A (based on his or her baseline

Figure 3: The Principles of Propensity Score Matching in Patients with Multiple Sclerosis



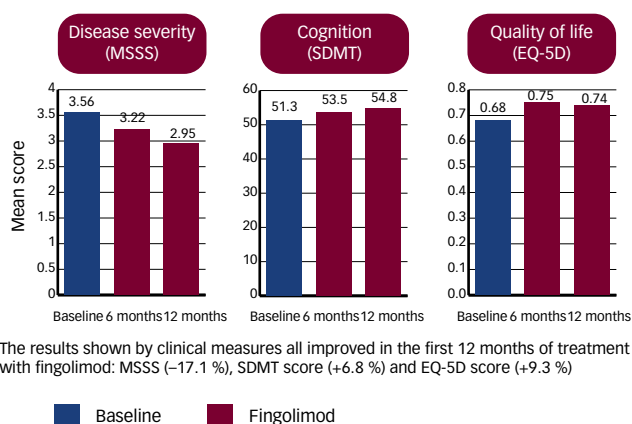
EDSS = Expanded Disability Status Scale; GA = glatiramer acetate; IFN = interferon. Source: Rosenbaum and Rubin 1983.⁶

Figure 4: Propensity Matching to Identify Matched Populations in Multiple Sclerosis



Source: Rosenbaum and Rubin 1983.⁶

Figure 5: IMSE II – Changes in Clinical Measures after Fingolimod Treatment for at Least 12 Months



The results shown by clinical measures all improved in the first 12 months of treatment with fingolimod: MSSS (-17.1 %), SDMT score (+6.8 %) and EQ-5D score (+9.3 %)

EQ-5D = European Quality of Life-5 Dimensions questionnaire; MSSS = Multiple Sclerosis Severity Score; SDMT = symbol digit modalities test. Reproduced with permission from Nordin et al., 2013.¹⁴

characteristics) will be matched with a patient in group B who also has a propensity score of 0.41 of receiving treatment A (see Figure 3). On the basis of this probability, individuals with the same propensity score can then be matched (as perfect pairs) in the two groups to be compared, and subsequent outcome analysis (e.g. relapse rate, disability progression, DMD discontinuation) can be performed using powerful paired statistics.

By way of further explanation, in a randomised control trial, the probability of being treated with DMT A or B is fixed, often at 50:50, and thus the propensity score (probability of receiving treatment A) is 0.5, whereas in the real-world data the probability of treatment A or B are different, so the population that can be studied in propensity-matched analyses is the overlap that occurs between the treatment assignment probabilities (see Figure 4). One effect is that patients with unmatched scores, e.g. 0.01 or 0.99, whose treatment assignment is almost always either A or B, are removed from the subsequent analysis. In practice, this could often include patients with either very mild or very severe disease initially. Questions on the validity of such analysis are raised, and MSBase proposes that one important means of validation is the capacity to assess treatment comparison for which the outcome is known from a randomised clinical trial, and several of these projects are underway, delivering results largely concordant with clinical trials.^{7,9}

An MS registry programme in clinical practice will improve the management of this chronic disease, especially where a large number of patients are seen. By participating in a registry, benchmarking and decision support can be enhanced, adding value to a clinical practice. Furthermore, such registries and databases are the only feasible way to generate long-term safety and efficacy data that are clinically meaningful to patients and may facilitate personalised medicine especially if combined with biomarker and MRI data. Importantly, real-world data can be benchmarked against randomised clinical trial results to demonstrate validity and improve the confidence of results generated from registries.

Collection of Fingolimod Real-World Evidence and Use in Clinical Practice and Research

A number of tools developed over the last few years can be used to investigate real-world data. It is increasingly important to collect pharmaco-economic data, which can be facilitated using claims databases, observational studies and population-based registries. Claims databases represent an easy way to obtain data as the DMTs and services are paid for, and the data collected can be interrogated to establish the effect of DMTs on resource utilisation in the clinical setting and medical costs. PharMetrics Plus™ is one of the most important US claims databases and processes the billable interactions between patients and healthcare providers.^{10,11} This database has information on over 200,000 patients with MS claims diagnosed between 2006 and 2013.¹¹

Considerable data have been collected on fingolimod, the first oral MS therapy. The other oral therapies were not available at the time these data were collected. PharMetrics Plus has data on a matched cohort of patients who switched to fingolimod or glatiramer acetate (GA) from interferons (IFNs). Those switched to fingolimod had a reduced claims-based relapse rate compared with GA over 12 months (62 %; $p=0.0013$).¹¹ A number of other parameters can be investigated, such as resource utilisation, allowing the quality of life of patients for the money invested to be determined. In this patient group, inpatient visits were reduced by 71 % ($p<0.01$) in patients switched to fingolimod compared with before switch. A similar reduction (67 %; $p<0.001$) was observed for the number of days corticosteroids were supplied.¹² These data provide invaluable insights into the effect of a DMT.

Another approach involves population-based registries, which have been developed especially in Scandinavia and can reveal the long-term impact of a DMT on clinical, safety, resource-use and socioeconomic outcomes. In the Swedish MS registry set up in 2002 (>60 healthcare units), information on approximately 12,000 patients includes treatment

Table 1: Recommended Magnetic Resonance Imaging Brain Protocol in Adult Multiple Sclerosis Patients – Consortium of Multiple Sclerosis Centers Guidelines

Sequence	Diagnostic Scan for CIS	MS Baseline or Follow-up Scan	Comment
1. 3 plane (or other) scout	Recommended	Recommended	Set-up axial sections through subcallosal line ^a
2. Sagittal fast FLAIR	Recommended	Optional	Sagittal FLAIR sensitive to early MS pathology, such as in corpus callosum
3. Axial FSE PD/T2	Recommended	Recommended	TE1 minimum (e.g. ≤30 ms). TE2 (usually ≥80 ms). PD series sensitive to infratentorial lesions that may be missed by FLAIR series
4. Axial fast FLAIR	Recommended	Recommended	Sensitive to white matter lesions and especially juxtacortical-cortical lesions
5. Axial pre-gadolinium T1	Optional	Optional	Considered routine for most neuroimaging studies
6. 3D T1	Optional	Optional	Some centres use this for atrophy measures
7. Axial gadolinium enhanced T1	Recommended	Optional	Standard dose of 0.1 mmol/kg injected over 30 s; scan starting minimum 5 min after start of injection

CIS = clinically isolated syndrome; CMSC = Consortium of Multiple Sclerosis Centers; FLAIR = T2-fast attenuated inversion recovery; FSE = fast spin-echo (or turbo spin-echo); PD = proton density-weighted (long TR, short TE sequence); T2 = T2-weighted (long TR, long TE sequence); T1 = T1-weighted (short TR, short TE sequence). Section thickness for sequences 3–6 is ≤3 mm. ^aThe subcallosal line joins the undersurface of the front (rostrum) and back (splenium) of the corpus callosum. Reproduced with permission from Simon et al., 2006.¹⁹

with DMTs, demographics, EDSS, relapses and MRI.¹³ It can also be linked to other population-based registries. Immunomodulation and Multiple Sclerosis Epidemiology Study of Fingolimod (IMSE) II is one of the five post-marketing Swedish surveillance studies monitoring DMTs, investigating the long-term effectiveness and safety of fingolimod in 806 patients from 44 centres.¹⁴ The majority were relapsing–remitting MS (RRMS) patients most frequently treated with natalizumab or IFNs (mean duration 11.5 months). Patients switched to fingolimod had a reduced annualised relapse rate (0.15 with fingolimod versus 0.39) and a more stable EDSS score over 12 months compared with before switch. In addition, more innovative parameters were investigated such as disease severity (Multiple Sclerosis Severity Score [MSSS]), cognition (Symbol Digit Modalities Test [SDMT]) and quality of life (European Quality of Life 5 dimensions questionnaire [EQ-5D]) and over 12 months all of these clinical measures improved with fingolimod (MSSS, –17.1 %; SDMT, +6.8 %; EQ-5D score +8.3 %) (see Figure 5). Eventually, long-term data will be available.

Finally, observational studies can be conducted to investigate the impact of one DMT on multiple patient- and clinical-reported outcomes. An example from Germany is Post-Authorization Non-interventional German sAFety study of GilEnyA in RRMS patients (PANGAEA), a prospective, observational, 5-year registry study of patients treated with fingolimod.¹⁵

Over a number of years, software tools have been developed to provide and collect data. For example, the Multiple Sclerosis Documentation System (MSDS3D) – designed as a database application for the documentation of MS patients – which is used for PANGAEA.¹⁶ It can interact with various patient-monitoring systems, such as clinical documentation, and specific treatment management for the more recently licensed DMTs, such as alemtuzumab, natalizumab, fingolimod, dimethyl fumarate or teriflunimide, and can be linked to different registries and for research projects. Its architecture is innovative, with the data stored in the MSDS3D cloud. The MS nurse can add most patient information via an iPad with only specific clinical aspects, such as adverse events and patient-related outcomes, reported by the physician. There is also the facility to present case reports for discussion.

The real-world evidence programme has demonstrated the benefits of fingolimod in regular clinical practice. Considerable claims outcome data has been collected for a large population of patients but to date imaging measures of brain tissue damage are missing and such data would undoubtedly strengthen real-world evidence.

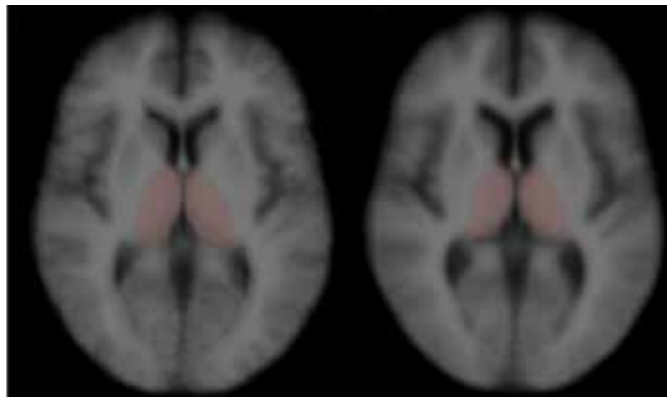
Evolving the Evidence Base – Importance of Collecting Imaging Measures

Although imaging data is routinely collected to assist in diagnosing MS and following disease progression, to date such data have not been systemically included in real-world databases. Conventional MRI sequences include, among others, T2-fast attenuated inversion recovery (FLAIR), proton density and T2-weighted imaging (T2/PD-weighted imaging). Over the last 10 years, improvement in scanner technology has made these techniques more sophisticated for performing quantitative analysis.¹⁷ FLAIR and T2/PD-weighted imaging are important sequences in the clinical routine for recognition of hyperintense lesions (T2 volume and number). Also use of pre- and post-contrast T1 spin echo scans help identify hypointense T1 and gadolinium (Gd)-enhancing lesions, respectively. Gd-enhancing lesions represent acute inflammatory activity indicative of the breakdown of the blood–brain barrier.¹⁸ About a third of these lesions evolve over time into chronic hypointense T1 lesions or black holes that are related to neurodegeneration in MS. Three-dimensional (3D) T1-weighted images are specific for volumetric assessments and better morphology tissue detection and with acceleration acquisition techniques, high-quality images can be obtained in only a few minutes. Such images allow tissue segmentation of grey matter, white matter and cerebrospinal fluid compartments. Historically these techniques were performed in academic centres but are now increasingly available in community practice. However, it is not necessary to acquire these sophisticated sequences to perform quantitative analysis in MS patients.

A number of recommendations for brain MRI protocols are available for adults with MS, for example, the Consortium of Multiple Sclerosis Centers (CMSC) guidelines (see Table 1).¹⁹ New improved scanners produce 3D FLAIR sequences that can give different orientation of the slices including sagittal, coronal and axial, without acquisition of each plane separately. It is important to note that these images can be obtained on 1.5 or 3 Tesla scanners and the recommended slice thickness is 3 mm without gap. In the past this was problematic due to the time involved, but in the last few years it has become possible to obtain this recommended protocol in less than 30 minutes, especially using acceleration acquisition techniques, such as parallel imaging.

An important question to address is whether collection of MRI data is feasible in clinical routine. There are a number of general requirements for image processing that need to be considered:

Figure 6: Comparison of 26 Healthy Controls Age- and Sex-matched – (Left) to 98 RRMS patients; (Right) on Averaged 3D T1-weighted Magnetic Resonance Images



RRMS = Relapsing-remitting multiple sclerosis Individual patient images were co-registered into the Montreal Neurological Institute 152 space for visualisation. Overlaid in red is the average thalamus volume for each group as calculated by FreeSurfer. The thalamus volume is 13.5 % in the multiple sclerosis (MS) group. The third ventricle is also clearly enlarged in the MS patients. Reproduced with permission from Minagar et al., 2013.²³

- Provide the technological resources to perform accurate, reproducible analyses of medical imaging data from which sound routine and scientific conclusions can be drawn.
- Maintain data integrity and confidentiality and ensure its future availability as techniques are constantly evolving.
- Make the workflows involved as efficient as possible – many techniques are becoming fully automated, being suitable for routine clinical use and make virtually no errors:
 - receive scans;
 - store scans;
 - analyse scans;
 - ensure scan analysis quality;
 - store analyses for future use;
 - maintain data integrity; and
 - report data within a few hours to the physician.

In many centres, picture archiving and communication systems (PACS) store the MRI scans, which is important for allowing future data analysis. Extremely powerful systems, able to perform analyses on a large amount of data, are now relatively inexpensive. However, as well as with the computational storage system, it is crucial that as much of the analysis as possible is automated ensuring consistent workflow. In the future the role of the operator in performing directly analyses will substantially change. The human interaction will become only necessary for processes requiring quality control checks to determine whether the analysis is acceptable to be used on an individual patient basis.²⁰ Quality checks are undertaken at various stages during the process to reduce or remove scanner error, diminish variability, prevent unanalysable scans and increase the power to detect effects earlier or in smaller groups.

The impact of pre-processing techniques is playing a major role in avoiding scanner or motion artefact errors. One of the most important tools in imaging pre-processing is non-uniformity correction, which removes low spatial frequency background perturbations via iterative, non-parametric bias field estimation and performs adjustment of the scan without changing its quality. Another major component is co-registration, which realigns images in 3D in order to correct for positioning errors

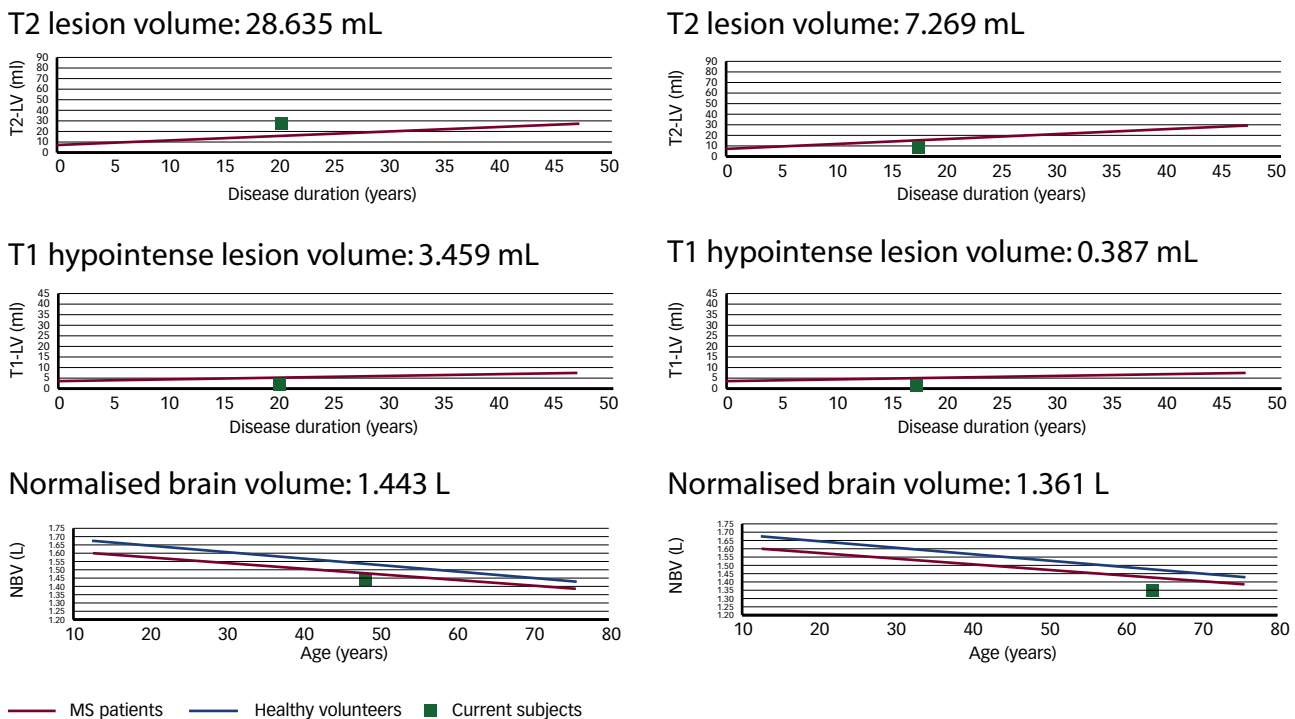
and/or patient movements. New lesions are sometimes reported in MS patients, while some of these represent positioning errors. Both these techniques occur automatically in a few seconds with virtually zero errors. Multidimensional combination of quantitative analysis is performed and although there are many different types, they can generally be divided into focal (Gd-enhancing lesions, T2 lesions, T1 black holes), macroscopic or visible global (atrophy, tissue-specific atrophy, regional atrophy) and microscopic or invisible global (diffusion-tensor, magnetisation transfer ratio, perfusion, spectroscopy, phase [iron]). Quantitative activity analysis of lesions can help to reduce the time of examining scans. If a new lesion appears on FLAIRs on longitudinal images, a new post-processing technique, the subtraction, can give a clear indication whether an apparently new lesion is real – such verification is invaluable as part of the clinical routine.

Many measures of brain atrophy are currently available and this is an extremely important outcome both for clinical trials and in predicting disability. However, the situation is complicated because during the disease course, the importance of some of these measures is changing. For example, in very early disease (clinically isolated syndrome [CIS] or early MS) thalamus and other grey matter structures evolve more rapidly compared with normal controls.²¹ However, later in the disease process, cortical atrophy is more involved. By contrast, lateral ventricular volume enlargement is independent of disease course and unfolds linearly across time. It is very predictive of clinically definite MS, disability development in CIS patients or secondary progressive MS (SPMS).²² However, to obtain a more comprehensive picture, other volumetric components, such as corpus callosum atrophy of the white matter, have to be also considered.

To address whether these measures can be used on an individual basis, 26 age- and sex-matched healthy controls were compared with 98 RRMS patients on 3D-T1-WI (see Figure 6). Although the images look relatively the same, the thalamus volume is 13.5 % lower in the RRMS group, illustrating that it is almost impossible to visually define the changes, necessitating the move to quantitative tools.²³

The Buffalo Neuroimaging Analysis Center at the University at Buffalo has extensive experience collaborating with long-term MRI databases. A 10-year collaboration with Charles University, Prague on the collection of data from the Avonex-Steroid-Azathioprine clinical trial.²⁴⁻³⁰ yielded important insights into understanding the natural history of brain atrophy development under DMTs. This is one of few studies that used the same scanner without hardware or software changes to acquire serial yearly scans over 10 consecutive years. The 5-year data showed a good correlation between disability and whole brain volume loss, cortical atrophy and lateral ventricle volume enlargement but no correlation with T2 lesion volume was found.²⁹

Furthermore, this centre has also collected MRI data from more than 2,500 MS patients and 400 healthy controls on the same 3 Tesla and 1.5 Tesla scanners over 8 years using standardised MRI clinical routine protocols.³¹⁻³³ MRI scans are stored daily in a centralised database system and fully automated computational systems run the analyses, which are overwritten and checked by the operators. All these data are available to physicians. There are multiple examples that can illustrate dynamic changes of lesion volume accumulation and brain atrophy progression, as outcomes on an individual patient basis from that database. Figure 7A shows the patient that had approximately 30 ml of T2 lesion volume after 20 years of disease duration, which

Figure 7: Lesion Volume and Brain Atrophy as Outcomes on an Individual Patient Basis

Left panel: Patient with typical lesion volume accumulation after 20 years disease duration. Right panel: Patient with lower T2- and T1-lesion volumes than expected for the disease duration. Note magnetic resonance imaging scan performed: 30 days from relapse or steroid treatment.

is representative of lesion accumulation over this time period in a typical MS patient. Based on a group of more than 1,000 matched MS patients and over 400 matched healthy controls, this patient fell 25 percentiles below the standard line of the MS patient and control age-related brain volumes, indicating more brain atrophy. This is expected since MS patients with more lesions tend to have increased brain atrophy. However, *Figure 7B* shows another example of an individual patient with lower T2- and T1-lesion volumes than expected for the disease duration, but with more advanced brain atrophy in terms of the matched MS and control groups. Such patients are more difficult to treat because it is believed that their disability progression is not so much dependent on lesions, but is due to rapid brain atrophy development. These real-world cases indicate that brain atrophy should be part of the clinical routine evaluation.

It is feasible to use the quantitative tools described in a clinical routine since the quality of MRI sequences is improving dramatically in last couple of years. The collection, storage and computation of the images are now more straightforward and the use of huge database systems together with automated analysis pipelines revolutionises the analysis. Furthermore, use of pre-processing techniques is making a dramatic difference in improving the quality of scans and facilitating reliable analysis.

Improvements in standard MRI protocols allow for quantitative data assessment in academic and community centres worldwide. In addition, PACS and other storage systems are available in most centres and allow easy MRI transfer and collection. Software improvements in MRI analysis are closing the gap towards application in clinical routine. It is clear that systematic MRI data collection on an individual patient basis is possible, but ideally should be further standardised. In the future, the focus should be on simplifying the

measures described to enhance their role in the follow-up of MS patients since improving follow-up will make a significant difference to patient management.

Real-world Evidence – Evolving Outcome Measures

Real-world evidence provides information about the effectiveness and long-term safety of a treatment in everyday clinical practice, but currently the systematic collection of MRI data is missing. To strengthen real-world evidence, the collection of clinical outcomes needs to evolve to include four core measures – relapses, EDSS, new lesion development and brain volume loss. Another outcome measure of great potential is tracking cognitive change over time, but the challenge of its implementation is even greater than that of routinely quantified MRI, due to its resource intensity. Large democratic data collection platforms, such as MSBase, make it easy for practising neurologists to form collaborations for collective outcome analysis and knowledge exchange.

Data from randomised, controlled trials can be expanded and enhanced with real-world data. For example, the systematic collection of real-world evidence has confirmed the effectiveness and safety of fingolimod in the real-world setting. First-hand real-world evidence is generated daily by clinical practices and a valuable contribution would be made to real-world MS research if data were routinely submitted to a local or international registry.

The authors are therefore confident that increased precision in monitoring will occur in the near future. Over the next few years, it is hoped that it will be possible to access other measures such as cognition and quantitative MRI, which together with clinical outcome measures, will facilitate improvements in treatment and treatment-failure definitions. ■

- European Database for Multiple Sclerosis (EDMUS), 2014. Available at: <http://www.edmus.org/en/proj/index.html> (accessed 30 June 2014).
- <http://www.neuroreg.se> (accessed 23 June 2014).
- Koch-Henriksen N, The Danish Multiple Sclerosis Registry: a 50-year follow-up, *Mult Scler*, 1999;5:293–6.
- Multiple sclerosis database (MBbase), 2014. Available at: <https://www.msbase.org/> (accessed 30 June 2014).
- <http://clinicaltrials.gov/ct/search?term=multiple+sclerosis> accessed June 2014.
- Rosenbaum PR, Rubin DB, The central role of the propensity score in observational studies for causal effects, *Biometrika*, 1983;70:41–55.
- Kalincik T, Spelman T, Trojano M, et al., Persistence on therapy and propensity matched outcome comparison of two subcutaneous interferon beta 1a dosages for multiple sclerosis, *PLoS One*, 2013;8:e63480.
- H Butzkueven, T Kalincik, V Jokubaitis, et al., Head-to-head comparisons of interferon beta and glatiramer acetate in clinical practice, *Mult Scler J*, 2014;20:927–927.
- T Spelman, N Bergvall, D Tomic, et al., Real-world comparative effectiveness of fingolimod and interferon/glatiramer therapies in a switch population using propensity-matched data from MSBase, *Mult Scler J*, 2013;19:514, P1096.
- IMS Health IMS Metrics PlusTM Scientifically proven and fit for purpose 2013 <http://myemail.constantcontact.com/IMS-PharMetrics-Plus-Scientifically-proven-and-fit-for-purpose.html?oid=1102157698477&aid=a8Gnc-fBspY> (accessed 30 June 2014).
- Bergval N, Makin C, Lahoz R, et al., Relapse rates in patients with multiple sclerosis switching from interferon to fingolimod or glatiramer acetate: a US claims database study, *PLoS One*, 2014a;6:e88472.
- Bergval N, Lahoz R, Reynolds T, Kom JR, Healthcare resource use and relapses with fingolimod versus natalizumab for treating multiple sclerosis: a retrospective US claims database analysis, *Curr Med Res Opin*, 2014b;30:1461–71.
- Westerlind H, Ramanujam R, Uvehag D, et al., Modest familial risks for multiple sclerosis: a registry-based study of the population of Sweden, *Brain*, 2014;137:770–78.
- Nordin N, Hillert J, Piehl F, et al., The immunomodulation and multiple sclerosis epidemiology (IMSE II) study: a Swedish nationwide pharmaco-epidemiological and genetic study focussed on long-term safety and efficacy of fingolimod (Gilenya®), ECTRIMS 2013 P975, *Multiple Sclerosis*, 2013;19:445–6.
- Ziemssen T, Diaz-Lorente, Fuchs A, Cornelissen C, 24-month interim results of PANGAEA: A 5-year registry study evaluating long-term safety, efficacy and pharmaco-economic data of German multiple sclerosis patients on fingolimod therapy, *AAN*, 2014;P3.152.
- Ziemssen T, Kempcke R, Eulitz M, et al., Multiple sclerosis documentation system (MSDS): moving from documentation to management of MS patients, *J Neural Transm*, 2013;120(Suppl. 1):S61–6.
- Zivadnov R, Stosic M, Cox JL, et al., The place of conventional MRI and newly emerging MRI techniques in monitoring different aspects of treatment outcome, *J Neurol*, 2008;255(Suppl 1):61–74.
- Hein J, van Waesberghe TM, van Walderveen MAA, et al., Patterns of lesion development in multiple sclerosis: Longitudinal observations with T1-weighted spin-eco and magnetization transfer MR, *AJNR Am J Neuroradiol*, 1998;19:675–83.
- Simon JH, Li D, Traboulsee A, et al., Standardized MR imaging protocol for multiple sclerosis: Consortium of MS Centers consensus guidelines, *AJNR Am J Neuroradiol*, 2006;27:455–61.
- Souplet JC, Lebrun C, Chanalet S, et al., Approaches to segment multiple-sclerosis lesions on conventional brain MRI, *Rev Neurol (Paris)*, 2009;165:7–14.
- Zivadnov R, Havrdova E, Bergsland N, et al., Thalamic atrophy is associated with development of clinically definite multiple sclerosis, *Radiology*, 2013b;268:831–41.
- Pagani E, Rocca MA, Gallo A, et al., Regional brain atrophy evolves differently in patients with multiple sclerosis according to clinical phenotype, *AJNR Am J Neuroradiol*, 2005;26:341–46.
- Minagar A, Barnett MH, Benedict RH, et al., The thalamus and multiple sclerosis: modern views on pathologic, imaging, and clinical aspects, *Neurology*, 2013;80:210–19.
- Dolezal O, Dwyer MG, Horakova E, et al., Detection of cortical lesions is dependent on choice of slice thickness in patients with multiple sclerosis, *Int Rev Neurobiol*, 2007;79:475–89.
- Havrdova E, Zivadnov R, Krasensky J, et al., Randomized study of interferon beta-1a, low-dose azathioprine, and low-dose corticosteroids in multiple sclerosis, *Mult Scler*, 2009;15:965–78.
- Horakova D, Dwyer MG, Havrdova E, et al., Gray matter atrophy and disability progression in patients with early relapsing-remitting multiple sclerosis: a 5-year longitudinal study, *J Neurol Sci*, 2009;282:112–119.
- Horáková D, Kyr M, Havrdová E, et al., Apolipoprotein E ε4-positive multiple sclerosis patients develop more gray-matter and whole-brain atrophy: a 15-year disease history model based on a 4-year longitudinal study, *FoI Biol (Praha)*, 2010;56:242–51.
- Bergsland N, Horakova D, Dwyer MG, et al., Subcortical and cortical gray matter atrophy in a large sample of patients with clinically isolated syndrome and early relapsing-remitting multiple sclerosis, *AJNR Am J Neuroradiol*, 2012;33:1573–8.
- Zivadnov R, Tekwe C, Bergsland N, et al., Bimonthly evolution of cortical atrophy in early relapsing-remitting multiple sclerosis over 2 Years: A longitudinal study, *Mult Scler Int*, 2013;2013:231345.
- Zivadnov R, Bergsland N, Dolezal O, et al., Evolution of cortical and thalamus atrophy and disability progression in early relapsing-remitting MS during 5 years, *Am J Neuroradiol*, 2013c;34:1931–9.
- Zivadnov R, Weinstock-Guttman B, Hashmi K, et al., Smoking is associated with increased lesion volumes and brain atrophy in multiple sclerosis, *Neurology*, 2009;73:504–10.
- Zivadnov R, Marr K, Cutter G, et al., Prevalence, sensitivity, and specificity of chronic cerebrospinal venous insufficiency in MS, *Neurology*, 2011;77:138–44.
- Dwyer MG, Bergsland N, Zivadnov R, Improved longitudinal gray and white matter atrophy assessment via application of a 4-dimensional hidden Markov random field model, *Neuroimage*, 2014;90:207–17.