Treatment strategies in multiple sclerosis: Current and future practice considerations

Dr Paul Gallagher MBChB, BSc (MedSci)Hons, MRCP (Neurology)

A thesis submitted for the award of the degree of Doctor of Medicine (MD)

College of Medical, Veterinary and Life Sciences

University of Glasgow

June 2019

Abstract

This thesis focuses on clinically-relevant questions regarding the care of patients with relapsing remitting multiple sclerosis (RRMS), including appraisal of the current relevant literature regarding the disease itself and therapeutic approaches. An understanding of the nature of MS is detailed in a review of its characteristics and natural history in the pre-treatment era, to establish a baseline upon which the utility of disease modifying therapies (DMTs) in the modern era can be evaluated. A review of the history of therapeutics in MS leads into an appraisal of trials in the modern era, including potential and actual pitfalls in their design and interpretation. A comprehensive tabulation of the pivotal placebo-controlled and head-to-head DMT trials is included and allows an appreciation of the complexity in evaluating the evidence-base for the therapeutic options currently available. Inter-trial comparisons are often made despite differing patient populations, methods and outcome analyses and these are discussed. The potential for new therapeutic options is also considered in light of emerging evidence for novel treatment approaches.

The original work in this thesis stems from practical clinical dilemmas in the management of patients with RRMS and uses real-world observational data to address them. The available evidence-base supports the early initiation of DMTs in RRMS on short-term efficacy grounds, but there remains controversy regarding the timing and aggressiveness of therapeutic intervention, particularly with higher risk treatments, and uncertainty regarding the impact on long-term outcomes. We have studied the safety and efficacy of the oral DMTs dimethyl fumarate and fingolimod in our own centre, and describe their safety and efficacy in real clinical practice. The availability of oral DMTs in MS therapeutics was a watershed event and our cohorts exemplify the desire of patients to switch from injectable therapies when alternatives became available. Whilst our data support the efficacy of these therapies, their side effects remain limiting for a proportion of patients, and this was higher than expected from previous trials. Additionally, we have been able to describe potential risk factors for lymphopaenia with dimethyl fumarate, and identify cases where this can persist despite drug discontinuation, which is relevant given its (post-licensing) association with progressive multifocal leukoencephalopathy (PML).

In identifying RRMS patients with similar disease profiles within two centres in Scotland who were often treated differently, we have provided evidence of significant variation in practice in a close geographical area. Additionally, whilst outcome comparisons were hampered by methodological issues, there were significant reductions in some disease measures when DMTs were started sooner rather than later in statistically-matched cohorts. These cohorts now provide the opportunity for a prospective study to compare more detailed long-term outcomes in patients treated or not in the early stages of their disease. Lastly, a unique dataset is analysed to evaluate the safety and efficacy of switching from one powerful anti-inflammatory immune therapy to another, namely natalizumab to alemtuzumab, in highly active RRMS. This is an increasingly used strategy since the worldwide licensing of alemtuzumab, despite little evidence upon which to base the approach. The use of alemtuzumab in the UK and Ireland for many years before this licensing, as a result of its development in Cambridge, provided a multicentre cohort of patients with longitudinal follow-up unavailable elsewhere, and we have demonstrated that this sequencing appears safe and effective. Additionally, the management of the switch between these two treatments is a dilemma in itself and we present data to support a direct switch and avoiding prolonged delay beyond excluding the possibility of (subclinical) infections.

The studies presented here are therefore of real utility to the practising clinician in MS therapeutics.

Chapter Summary

Introduction

This describes the epidemiology, pathogenesis and clinical features of MS but primarily focuses on natural history studies to contextualise the impact of DMTs in RRMS, as shown in the pivotal clinical trials which are also described.

Chapter 1: DMT initiation and escalation in RRMS

This chapter reviews the current literature on DMT initiation and escalation in RRMS followed by a summary of real-world studies from our own centre on recently-introduced oral DMTs that are used in practice as both initiation and escalation agents.

Chapter 2: Multiple sclerosis Outcome Determination Evaluating Real Differences After TimE (MODERATE)

This chapter is a study of DMT initiation in RRMS in Scotland and highlights the variability and potential impact on patients of differing approaches.

Chapter 3: Alemtuzumab after Natalizumab Switch in Evolving Rapidly Severe Multiple Sclerosis (ANSWERS MS)

This chapter evaluates the safety and efficacy of sequencing highly-efficacious DMTs with profound immunomodulatory effects, based on unique multicentre longitudinal data from the UK and Ireland.

Conclusions

This thesis explores current treatment strategies in RRMS at local, national and international levels and their potential impact on patients now and in the future. Data presented here are part of the growing body of work recognising the necessity and utility of real-world observational studies and includes safety and efficacy data on commonly used DMTs in our centre as well as less common sequencing strategies with highly efficacious treatments using multicentre data. Sequencing highly effective DMTs as a treatment strategy in RRMS has a very limited evidence-base but this work provides reassuring data on both safety and

efficacy fronts with regards to the use of alemtuzumab after natalizumab, as well as favouring a shorter switch period where possible. Additionally, the variability of DMT use within Scotland, and the potential impact this can have, is demonstrated and provides a basis for further work. The utility of multicentre collaboration is inherent in the two main studies, MODERATE and ANSWERS MS, and serves a model for future collaborative studies to enhance our understanding of the role of current and emerging therapies in improving the lives of people with MS.

Acknowledgements

I would like to thank a number of contributors to the work in this thesis.

Firstly, I could not have hoped for a more approachable, enthusiastic, thoughtful and generous supervisor than James Overell. He has been instrumental in developing my interest in multiple sclerosis and clinical research, which I hope to continue throughout my career, and has become a valued colleague and friend over this time.

For the work on dimethyl fumarate use in our centre, I am indebted to Frederick Winslow (Medical student), Sheena Murdoch (Neurology Registrar) and Niall MacDougall (Neurology Consultant) for their assistance in data collection and analysis.

For the work on fingolimod use in our centre, I am very grateful to Kieran Fitzpatrick (Pharmacist) for the data collection and analysis which he shared. Sarah-Jane Martin (Neurology Registrar) assisted in the collection and analysis of data for the evaluation of fingolimod-associated macular oedema and has provided much support (and caffeine!) throughout this process.

The MODERATE study was only possible with collaboration from our colleagues in Aberdeen and I am extremely grateful to Angus Macleod (Neurology Consultant) for his valuable input and oversight in this project.

The ANSWERS MS study was a real group effort of clinicians across the UK & Ireland and I thank all involved from each centre for this and for putting up with my endless emailing.

Mario Hair (Statistician) provided invaluable support for the analysis of data in both the MODERATE and ANSWERS MS studies. I wish to thank him particularly for the work he has done, always with a patient, thorough and enthusiastic approach.

Finally, I wish to thank my family for their support throughout all of this. My wife (Felicity) and children (Anabelle and Louisa) have had to do without me on many occasions whilst I worked on these studies and the write-up or undertook locum work to allow me to do this as my day job for over 2 years. At times this has been a difficult process but I am so grateful for the support and will

appreciate this always. My parents have been the bedrock of belief in myself throughout my life and have encouraged and supported me at every stage, from (mum) driving me around to photograph smoking chimneys for a school geography project to (dad) getting me out for a walk when worried about exams and now (both) babysitting my children. I am forever grateful.

Author's declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where states otherwise by reference or acknowledgment, the work presented is entirely my own.

Paul Gallagher

June 2019

Contents

Publications arising from this work

Gallagher P, McGuigan C, Nicholas R, Hobart J, Ford H, Petheram K, Sharrack B, Robertson N, Mazibrada G, Scolding N, Wilson M, Pearson O, Jones J, Overell J. Alemtuzumab after Natalizumab SWitch in Evolving Rapidly Severe Multiple Sclerosis (ANSWERS MS): long-term UK & Ireland experience. *Multiple Sclerosis Journal* 2018; **24**(2) Suppl:8-120 [*Abstract*]

Gallagher P.J., McGuigan C., Nicholas R., Hobart J., Ford H., Petheram K., Sharrack B., Robertson N., Mazibrada G., Scolding N., Wilson M., Pearson O., Silber E., Rog D., Turner B., Harikrishnan S., Duddy M., Gray O., Jones J., Overell J. Alemtuzumab after Natalizumab SWitch in Evolving Rapidly Severe Multiple Sclerosis (ANSWERS MS). *Multiple Sclerosis Journal* 2017; **23**(3) Suppl 1:357-358 [*Abstract*]

Gallagher P.J., Hair M., Macleod A., Overell J. Multiple sclerosis Outcome Determination Evaluating Real differences After TimE (MODERATE): Treatment variations and outcomes in Scotland. *Multiple Sclerosis Journal* 2017; 23(3) Suppl 1:876-877 [*Abstract*]

Winslow F., **Gallagher P.J.**, Macdougall N.J., Overell J.R. Lymphocyte recovery after discontinuation of dimethyl fumarate due to lymphopaenia. *Multiple Sclerosis* 2016; **22** Suppl 3:793-794 [*Abstract*]

Gallagher P.J., Murdoch S., Macdougall N.J., Webb S., Overell J.R. 1st year of experience with dimethyl fumarate in the largest prescribing centre in Europe. *Multiple Sclerosis* 2015; **23**(11) Suppl 1:763-76 [*Abstract*]

MacDougall N.J., **Gallagher P.**, Murdoch S., Murray L., Webb S., Overell J.R. Persistent lymphopenia after routine clinical use of dimethyl fumarate for relapsing remitting multiple sclerosis. *Multiple Sclerosis* 2015; **23**(11) Suppl 1:300 [*Abstract*]

A platform presentation of ANSWERS MS (Chapter 3) at the international European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS) was part of the main Scientific Sessions on 12th October 2018 in Berlin, Germany.

List of Tables

List of Figures

Abbreviations

ADL	Activities of Daily Living
AE	Adverse Event
ALC	Absolute Lymphocyte Count
ALT	Alanine Transaminase
ARR	Annualised Relapse Rate
AST	Aspartate Transaminase
AUC	Area Under the Curve
AZA	Azathioprine
BP	Blood Pressure
CDMS	Clinically Definite Multiple Sclerosis
CIS	Clinically Isolated Syndrome
CIs	Confidence Intervals
CMV	Cytomegalovirus
CNS	Central Nervous System
COWAT	Controlled Word Association Test
CVLT	California Verb Learning Test
D-KEFS	Delis-Kaplan Executive Function System
DMF	Dimethyl Fumarate
DSS	Disability Severity Scale
EAE	Experimental Autoimmune Encephalitis
ECG	Electrocardiogram
EDSS	Expanded Disability Status Scale
EMA	European Medicines Agency
FAME	Fingolimod Associated Macular oEdema
FS	Functional System
GA	Glatiramer Acetate
Gd+	Gadolinium-enhancing (MRI lesion)
GI	Gastrointestinal
Hb	Haemoglobin
HSCT	Haematopoietic Stem Cell Transplant
HTN	Hypertension
IFN	Interferon
IMD	Index of Multiple Deprivation
ITP	Idiopathic Thrombocytopaenic Purpura
ITT	Intention To Treat
IVIgs	Intravenous Immunoglobulins
JCV	John Cunningham Virus

LAQ	Laquinimod
LFTs	Liver Function Tests
LSOA	Lower Layer Super Output Area
MRI	Magnetic Resonance Imaging
MS	Multiple Sclerosis
MSFC	MS Functional Composite
MTX	Methotrexate
Nabs	Neutralising Antibodies
NEDA	No Evidence of Disease Activity
NS	Not significant
NSAID	Non-steroidal Anti-inflammatory Drug
OR	Odds Ratio
PASAT	Paced Auditory Serial Addition Test
PBO	Placebo
Plts	Platelets
PML	Progressive Multifocal Leukoencephalopathy
PPMS	Primary Progressive Multiple Sclerosis
PROMs	Patient Reported Outcome Measures
PS	Propensity Score
PSM	Propensity Score Matching
PTE	Pulmonary Thromboembolism
RCT	Randomised Controlled Trial
RES MS	Rapidly Evolving Severe Multiple Sclerosis
RIS	Radiologically Isolated Syndrome
RRMS	Relapsing Remitting Multiple Sclerosis
RTA	Road Traffic Accident
Rx	Treatment
SAD	Sustained Accumulation of Disability
SAE	Serious Adverse Event
SAP MS	Single Attack Progressive Multiple Sclerosis
SDMT	Symbol Digit Modalities Test
SE	Side Effect
SMC	Scottish Medicines Consortium
SMSR	Scottish Multiple Sclerosis Register
SPMS	Secondary Progressive Multiple Sclerosis
SRD	Sustained Reduction in Disability
Sx	Symptoms
T25FW	Timed 25-Foot Walk

TB	Tuberculosis
TIW	Three per week
ULN	Upper Limit of Normal
USS	Ultrasound Scan
WCC	White Cell Count

Introduction

This thesis pertains to the clinical management of Relapsing Remitting Multiple Sclerosis (RRMS), specifically the use of available Disease Modifying Treatments (DMTs). Multiple Sclerosis (MS) is the most common disease of the central nervous system in young adults and affects around 2 million people worldwide[1]. The World Health Organisation (WHO) and the Multiple Sclerosis International Federation (MSIF) estimate the global prevalence of MS at 30 per 100 000 persons[2], but this belies wide regional variations. There appear to be significant differences between continents in the distribution of MS worldwide[3]. The highest prevalence is found in Western Europe and North America, with lower rates in Eastern Europe, the Balkans and Australasia and the lowest rates in Asia, the Middle East and Africa. A meta-analysis of available European studies has described its epidemiology of MS in detail[4]. The total estimated prevalence rate of MS in Europe for the past 30 years is 83/100000 population, with a female:male ratio of around two. Prevalence rates are higher for women in all countries studied and the 35-64 years age group had the highest rates overall. Prevalence rates varied significantly, with the highest recorded in the north of Ireland (Donegal) at 216 cases/100000 (282 in women) and the lowest in Albania, 11 cases/100000. Similarly, the female:male ratio varied between 1.1 (Albania and Cyprus) to 3.4 (Donegal, Ireland). The mean European annual incidence is estimated at 4.3 cases/100000, with the highest incidence in Scotland (12 cases per 100 00 per year) but with highs of 11.6, 8.7 and 6.8 cases per 100 000 per year in Finland, Norway and Sardinia also.

No single cause of MS has been identified and no pathognomic finding exists in vivo. Whilst both genetic and environmental factors have been postulated, it seems likely that multiple sclerosis results from interplay between these. The relative contribution of genetic and environmental factors on an individual basis is likely on a spectrum from predominantly genetic to predominantly environmental and, given the lack of homogeneity of the condition, the possibility that multiple sclerosis represents the common end-point of a number of pathological processes remains. The standard statement that MS is a T-cell-mediated autoimmune disorder resulting in focal demyelination of CNS white matter causing neurological dysfunction has since been challenged by

therapeutic responses to targeted B-cell therapies, the presence of diffuse white matter inflammation, involvement of grey matter and the meninges, axonal loss in the absence of demyelination and early neurodegeneration raising the possibility of the inflammatory response being a secondary rather than primary insult. The extent to which a failure of normal remyelination, rather than demyelination, contributes to the process is also uncertain.

The most pertinent debate regarding the pathogenesis of MS is that between the 'outside-in' and 'inside-out' hypotheses. The 'outside-in' hypothesis has most widespread historical support and states that MS results from a primary dysfunction of the (peripheral) immune system with subsequent autoimmune targeting of the CNS causing damage. Indeed, the most widely studied animal model from which the majority of immunopathological mechanisms in human MS are surmised, Experimental Autoimmune Encephalitis (EAE), is based wholly on this hypothesis through inoculating mice with abnormal immune cells. Similarly, all current licensed disease-modifying therapies MS target the peripheral immune system primarily, albeit some of these drugs may have neuroprotective benefits also. The alternative and more recent 'inside-out' paradigm hypothesises that the primary abnormality is within the CNS and that the immune response is a consequence of this, stimulated by antigenic intracellular breakdown products liberated by the (degenerative) process. Evidence for both hypotheses is available, but the fact that the primary location of MS pathogenesis remains obscure provides a sense of the nascent understanding of this condition despite extensive research. A major development in the understanding of MS pathology was proposed by Luchinetti et al.[5], stating that four distinct phenotypes of MS can be discerned immunopathologically and, further, that this is specific to an individual throughout life[6,7]. Challenges to their findings are made, but along with clinical evidence supporting differing therapeutic responses between these phenotypes[8], it is tempting to attribute both the variable clinical phenotype and treatment response seen in practice to such heterogeneous pathology. Notably, these 'immunopatterns' do not reliably distinguish between the clinical phenotypes of relapsing or progressive MS[9].

Pathological diagnosis is the gold-standard in almost all areas of disease, but the inaccessibility of in vivo tissue in MS prevents this and results in a skew towards outlier cases that have, by definition, required a brain biopsy or post-mortem.

MS pathology includes a mixture of inflammation, demyelination and degeneration of variable extent, but which of these occurs first or should be the primary therapeutic target remains a matter of debate. In vivo brain biopsy samples are exclusively a product of clinical practice, usually to differentiate MS from other conditions, as undergoing a brain biopsy for entirely research purposes would be unethical given the associated risks. Post-mortem samples may be available in younger patients, with shorter disease duration, but more often represent an older population with longer disease duration.

The symptoms and signs of MS are a consequence of the neurological region affected by the process such that, essentially, any CNS-dependent function may be affected. In practice, however, the lesions of MS have a tendency to occur in certain CNS regions with predictable symptoms resulting. Notably, large swathes of brain regions can be affected without any symptomatic sequelae and it is well recognised that the number of symptomatic lesions usually belies a higher lesion load within non-eloquent regions of the CNS. As little as 5-20% of new lesions identified on brain MRI are symptomatic, with more recent studies suggesting the relationship is at the lower end of this range[10]. In contrast, lesions within the spinal cord are more likely to be symptomatic, given the higher concentration of neural pathways in a relatively small area, anatomically, but they too can be clinically silent. However, symptoms related to spinal cord lesions are a common first presentation of MS (transverse myelitis), as well as a preponderance for the optic nerves and brainstem. In fact, some clinicians consider a presentation of internuclear ophthalmoplegia (related to a lesion of the medial longitudinal fasciculus within the upper brainstem) in a young person as almost pathognomic of MS and, whilst other causes are possible of course, there is a rationale to this conclusion. Although grey matter involvement is clearly demonstrated as a consequence of MS, cortical dysfunction as a first symptom, such as seizure and cognitive presentations, are rare. Non-focal symptoms, such as fatigue or depression, may well predate focal symptoms but the consensus view remains that these are not considered of diagnostic value, albeit they can be more debilitating than localisable disability.

Understanding the natural history of a condition is crucial in both defining prognosis and assessing treatment effects. Studies from the pre-treatment era of RRMS are of course most relevant in understanding this. The inclusion of patients given treatment (which usually differs from that given today) skews any conclusions about its 'natural' history, both by the inclusion of patients with more severe disease, as well as any treatment effects. The natural history of MS is critical to this thesis, given the focus on DMTs, their impact on patients in terms of efficacy and safety and prognostic factors. We must be sure that treatments truly modify the disease and can only claim this once the natural history is understood, therefore this will be discussed first, followed by a summary of the pivotal trials for treatments deemed as disease-modifying.

The Natural History of MS

Longitudinal observational cohort studies of Multiple Sclerosis (MS) have been ongoing in a number of developed Western countries since the 1970s. These are based largely in Europe and North America and, whilst acknowledging methodological weaknesses, provide the most robust evidence available to understand the natural history of MS. Indeed, such cohorts are used as comparators for contemporary cohorts in order to measure treatment effects of subsequently-introduced therapies[11] as a pseudo-placebo arm. Of course, there are notable drawbacks in this approach, not least the changing definitions of MS diagnosis, advances in imaging and unforeseeable 'unknown unknowns' which may invalidate comparisons between historical and contemporary patients with a disease of uncertain aetiology. The purpose of these longitudinal cohort studies, however, was to describe the expected course of MS over time on a population-basis and extrapolate influential (ideally modifiable) factors which may suggest treatment targets: this will be the focus of this section.

The prolonged course of MS necessitates significant longitudinal follow-up to describe its natural history accurately. The most prominent and usually heralding event in MS is the relapse – an episode of neurological dysfunction which generally resolves to some extent. A relapse typically has a clear clinico-anatomical correlate, with the symptom predictably related to the area of focal CNS dysfunction, often reflected by MRI changes in the relevant region of the

brain or spinal cord. Yet, as outlined below, the symptoms occurring in a relapse are usually transient. Hence, these are not the main cause of disability in the longer term (although they may contribute to this). The progressive disability in the absence of relapses is the main contributor to loss of function for patients and has the greatest societal impact. This is not particularly associated with focal CNS dysfunction demonstrable on imaging but, rather, a diffuse, poorly-understood mechanism causing gradual loss of function over time, appearing neurodegenerative rather than the waxing and waning inflammatory nature of relapses. It is these two processes which define the course of multiple sclerosis in the individual and therefore is the focus of the natural history studies outlined here. The relationship between relapses and progressive disability remains contentious but is often the overriding clinical basis for DMT use in MS, predicated on the assumption that short-term proven reductions in relapse frequency and associated imaging correlates will translate into a reduction in long-term disability: there is evidence from follow-up studies of DMT-treated patients that this assumption holds true, as discussed later.

There are a number of highly-referenced longitudinal cohort studies from the modern era which are taken to describe the natural history of MS[12–19]. A selection of these are outlined for comparison in **Table 0-1**, adapted from Tremlett et al[20]. The variability of findings from these, particularly with regard to the timing of reaching disability milestones, likely reflects differences in methodology and diagnostic definitions over time. It may be that the findings represent a true divergence of MS natural history within global regions, however. The effect of treatments used in each cohort is purported to have no significant impact on the results of analyses of each individual cohort, but the apparent increasing duration in time to fixed disability (i.e. EDSS 6) in the newer studies (with more treated patients) is notable. However, there also appears to be an increasing duration of time to reach EDSS 6 in the PPMS cohorts in the newer studies, in the absence of any proven DMT benefits, suggesting the disease has become less disabling with time or, perhaps, reflecting the milder cases incorporated with the use of subsequent diagnostic guidelines. Similarly, there appears an increasing predominance of female patients in the more recent studies, which may impact upon outcomes, with males generally having a poorer prognosis.

The London Ontario cohort could be considered the 'purest' natural history study given the lack of any DMT use. That said, the broad-spectrum immunosuppressant treatments used in the Lyon cohort, for example, are largely no longer used because of lack of efficacy in comparison to the MS-specific DMTs now widely used. The assumption would therefore be that they do not alter the course of the disease in a meaningful way. The findings presented here are largely drawn from the London Ontario[21], British Columbia[15] and Rennes[17] cohorts given the pivotal nature of their reported findings. Across studies, there are notable differences in outcomes between patients seen prospectively from onset and those whose analyses include retrospective data. In the 'seen at onset' subgroup of the Ontario cohort, 84.7% were initially diagnosed with relapsing-remitting MS in comparison to 65.8% of the total cohort which included retrospective data. Similarly, just 28.6% of the prospectively followed subgroup had converted to SPMS by years 11-15 from onset in comparison to 57.6% of the total cohort. Some of this may reflect the smaller number of patients and shorter follow-up times for this sub-group but may also suggest more historical accuracy in prospective assessments, expected to be more indicative of the true situation.

Table 0-1: Natural history cohorts summary

Cohort Location	**London, Ontario, Canada**[21]	**Lyon, France**[13]	**Olmstead County, Minnesota, USA**[14]	**British Columbia, Canada**[15]	**Nova Scotia, Canada**[16]	**Rennes, France**[17]	**Lorraine, France**[18,19]
Setting	Population- and Clinic-based from single outpatient MS specialist clinic	Clinic-based from single outpatient MS specialist clinic (said to be 'representative of the general population')	Population-based from patients seen at Mayo clinic or Olmstead Community Hospital	Population-based from 4 outpatient MS specialist clinics serving the entire province	Clinic-based from Nova Scotia's only specialist MS centre	Clinic-based from regional MS centre in West France	Population-based from a single region but multiple centres
Years included	1972-1984	1976-1997	1991-2000	1980-2003	1979-2004	1976-2004	1996-2003
Diagnostic criteria	Probable or possible MS, Poser criteria[22]	Definite or probable MS, Poser criteria[22]	Definite MS, Poser criteria[22]		Definite MS, Poser[22] or McDonald[23] criteria	Definite MS, Poser criteria[22]	Definite or probable MS, Poser criteria[22]
Data collection	Retrospective and prospective	Retrospective from onset to first clinic visit, prospective thereafter	Retrospective	Retrospective from onset to first clinic visit, prospective thereafter			Retrospective and prospective

Follow-up duration	11.9 yrs from onset 9.8 yrs from 1st clinic visit (PPMS cohort only)	11 yrs from onset	19.3 yrs from onset	20.1 yrs from onset 8 yrs from first clinic visit	Not provided	12.8 yrs from onset	13.7 yrs from onset
N	**1099**	**1844**	**201**	**2837**	**1607**	**2054**	**2871**
% RRMS	66	85	94.5	88	83	78	87
% PPMS	33*	15	5.5	12	17	22	13
% Female	66	64	70	70	74	70	72
Mean age at onset (years)	30.5	31	31.2	30.6	Not provided	31.4	33
DMT treatments**	Nil other than steroids for relapse	49% treated (AZA>Cyclophos >IFN>MTX>Mitox)	25% treated (IFN/GA)	15.5% treated (IFN/GA)	Not stated (By 2014, 2240 patients were included with 57% treated with a DMT at any time since 1998[16])	56% treated for at least 6 months (IFN >Mitoxantrone> AZA>MTX> Cyclophos>GA)	Not stated (23% of PPMS cohort treated, mainly with cyclophosphamide[8])

Disability measurement	DSS 3, 6 and 8 collected retro- and prospectively	EDMUS scale collected retro- and prospectively	EDSS via exam or phone interview	(E)DSS recorded by neurologist at each clinic visit (>95% prospective)		EDSS recorded by neurologist at each clinic visit	
Definition of reaching EDSS 6	Not specified	All subsequent scores ≥6	Sustained at 6 months	Sustained at 6 months and all subsequent EDSS ≥6.0			Sustained at 6 months
Median time (years) to SPMS from onset of RRMS							
	-	19	-	19	-	16	20
Median time (years) to EDSS 6 from onset of MS [approx.]							
RRMS	15	23	-	30	35	21.7	24.5
PPMS	4.5	6	5	13	21	10	10

*Data unavailable for 0.9% of cases

** Azathioprine (AZA), Cyclophosphamide (Cyclophos), Interferons (IFN), Methotrexate (MTX), Mitoxantrone (Mitox), Glatiramer Acetate (GA), Disease Modifying Therapy (DMT)

Disease characteristics and predictors of outcome

In the Ontario cohort, detailed analysis of symptomatic onset was undertaken. There appeared an over-representation of younger patients (<20yrs) presenting with optic neuritis (22.9%) in comparison to older patients (>49yrs), in whom only 6.3% presented in this way; conversely, a higher proportion of older patients presented with insidious motor weakness (46.8%) in comparison to their younger counterparts (3.8%). Sensory symptoms at onset were more common in patients with younger onset but the differences were less striking (46.5% vs 31.9%) and age had even less relevance to ataxic presentations, occurring in 13.7%of those less than 20yrs old at the time of onset and 10.6% in those presenting beyond 49 years of age[21].

Factors associated with a more severe disease course were investigated in the Ontario cohort[12,24]. With regard to the early disease course, increased frequency of relapses within the first 2 years, a reduced interval between the 1st and 2nd relapse and a faster rate at which Disability Status Scale (DSS) 3 is reached were all found predictive of a more severe disease course over time and associated with an earlier need for walking aids (DSS 6). The Disability Status Scale (DSS) is a non-linear grading system initially proposed in the 1950s, ranging from 0 (no symptoms and normal neurological examination) to 10 (death due to MS); it largely relies on motor dysfunction to determine disability and predominantly lower limb, with DSS 6 reflecting the need to walk with a stick, for example, and was expanded in 1983 to include half-steps in the scale also (EDSS)[25,26]. In the untreated Ontario cohort, 66% had reached DSS 3 at 7 years after onset; in the cohort seen from onset, 62% had reached DSS 3 by 2 years. The frequency of relapses was higher in the first 2 years after disease onset in those who subsequently developed SPMS (within the follow-up period) in comparison to those who remained with a relapsing-remitting phenotype (mean 2.08 relapses vs. 1.8 in first year after onset in those who developed SPMS and those who did not, respectively; the values were 1.08 and 0.55 relapses per year in year 2)[12]. In addition, the following factors were statistically associated with a poorer outcome i.e. whether they significantly influenced the time to requiring a walking aid by the end of follow-up[24]:

- Older age at onset ($p<0.001$)

- Male sex (p=0.004)
- Ataxia at onset (p<0.001) [controlled for age at onset and gender]
- Insidious motor involvement at onset (p=0.01) [controlled for age at onset and gender]
- Cerebral, Cerebellar or brainstem involvement at last visit (all p<0.001) [controlled for age at onset, ataxia at onset and relapsing-remitting course]
- Higher relapse frequency in first 2 years (p<0.001) [controlled for age at onset, relapsing-remitting course, cerebellar and cerebral involvement at last visit]
- Shorter first inter-attack interval (p<0.001) [controlled for age at onset, relapsing-remitting course, cerebellar and cerebral involvement at last visit]
- Higher DSS at 2 and/or 5 years (p<0.001) [controlled for age at onset, relapsing-remitting course, cerebellar and cerebral involvement at last visit]

Optic neuritis at onset was associated with a better outcome i.e. less likely to require a walking aid during the follow-up period (p=0.03) and was controlled for age and gender.

Disease course phenotypes

A progressive decline in function, largely in the absence of relapses, can occur either from onset [Primary Progressive Multiple Sclerosis (PPMS)] or after a variable period of relapses and remissions [Secondary Progressive Multiple Sclerosis (SPMS)]. The proportion of patients presenting with PPMS appears to vary widely amongst cohorts, ranging from 9% in a French cohort in the 1960s[27] up to 37% in a Dutch cohort from the 1980s[28]. **Table 0-1** reflects this variability even in modern studies, however. Patients with PPMS tend to present with motor, rather than sensory or cerebellar deficits[29]. Additionally, again from the Ontario group, patients with PPMS developed disability quicker and died sooner than those who develop SPMS – notably, PPMS tends to present at an older age and more commonly in men, however. Axonal reserve is lower with age, potentially explaining the more rapid disability and both male sex and increased age are associated with increased risk of death in general, hence the

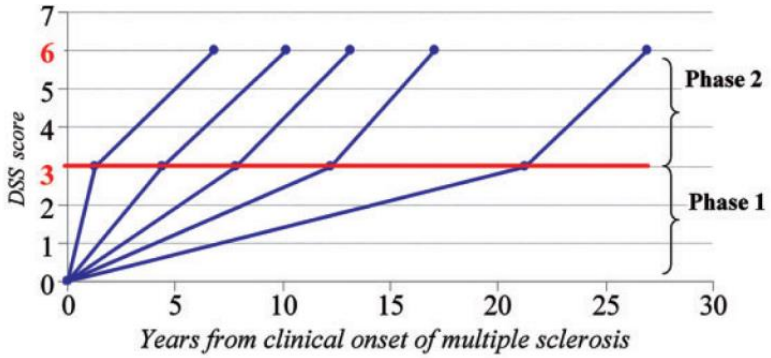
contribution of MS to these outcomes is not entirely clear, from this analysis at least. Relapses occurred in 27.8% of the Ontario group PPMS cohort: often mild and usually remitting, there was no significant difference in disability progression between those patients with early relapses and those without, thus the terms 'relapsing-progressive' and 'progressive-relapsing' are equivalent and were subsequently both included under the umbrella of PPMS[30]. Relapses can occur at any time in PPMS but tend to be within the first 10 years, as was the case for half of the Ontario PPMS cohort. That said, 10% had relapses after 20 years, up to a maximum of 39 years post-onset. The frequency and distribution of relapses was consistent however: most had a single event and were largely extra-spinal.

The conversion from RRMS to SPMS had occurred in approximately 80% of patients within 20 years from onset in the Ontario group but there remains uncertainty regarding the impact of relapses on the development of progression, albeit the progressive phase appears very uniform once it occurs[31,32]. Among patients with PPMS, SPMS and Single-Attack Progressive (SAP) MS [where a progressive course develops after a Clinically Isolated Syndrome (CIS)], there was no difference in time to disability milestones (DSS 6, 8 and 10). The authors conclude that the progressive course is independent of the relapses which occur before (or after) the onset of insidious disability progression. Notably, in SAP MS, the site of the CIS was not usually where progression began and the degree of recovery from the initial event was not related to longer-term outcomes. That is, 65% recovered from the initial event with no residual disability while 7% had fixed disability as a result, yet the subsequent rate of progression was the same for both groups[31]. Age may be important here, again, in understanding the impact on patients. Those with progression from the outset (PPMS) tend to become symptomatic at an older age (mean 38.5 years) than either SAP MS patients (33.3 years) or those who develop SPMS after a relapsing-remitting phase (29.8 years). Those with a single event preceding the progressive phase (SAP MS) had a mean latency of 7.6 years following the event before progression occurred; in contrast, the relapsing-remitting phase (mean 0.65 relapses/year[32]) took 10.3 years to develop progressive disease (SPMS).

Once the progressive course begins, the rate of progression appears the same, irrespective of the preceding phenotype[17,31]. This was suggested from the

Ontario group and delineated clearly by Leray et al. using the Rennes MS databse[17]. Analyses from this cohort and that from British Columbia are widely cited in the understanding of the relationship between relapses and disability. A few notable differences exist between the Rennes and British Columbia cohorts, although their results appear entirely compatible. Notably, whilst the Rennes cohort had an average of 12.8 years follow-up, 35% had reached DSS 6 in comparison to 28% of the British Columbia cohort where follow-up was beyond 20 years. This suggests a more severely affected cohort from Rennes, yet only 38% developed SPMS after a median of 16 years from onset in comparison to 55% in the British Columbia studies[33]. LeRay et al. found that disability progression from DSS3 to DSS6 was independent of the factors from diagnosis to DSS3, whether the course was relapsing-remitting or not as shown in **Figure 0-1**. This shows that, once DSS3 was reached, the mean time to DSS6 was 6-9 years whether the initial course was progressive or relapsing in 718 patients with MS (divided into 5 subgroups based on time to DSS3). Notably, given the higher proportion of treated patients in the Rennes cohort, the results were unchanged when the untreated cohort was analysed. Similar to the Ontario group, male gender, older age at onset, residual deficit after the first relapse and the number of relapses in the first 2 years were associated with a shorter time to irreversible disability (DSS 3) in the relapsing-remitting cohort.

Figure 0-1: Two-stage progression in MS (from LeRay et al.[17])



It is notable that the median age of onset of the progressive phase was similar in primary and secondary progressive cases – the authors suggest that MS may therefore be a chronic, neurodegenerative, age-related disease, unaffected by the initial course[17]. Whilst patients with PPMS progress at a faster rate, the age

at which they reach disability milestones is comparable to those with an initially relapsing course.

Relapses and progression

There remains controversy regarding the relationship between relapses and longer-term disability progression. Kremenchutsky et al.[31] state that the great majority of MS relapses are associated with full or partial recovery in their natural history study, whilst Lublin et al.[34] cite significant increases in disability measures caused by relapses as evidence of a substantial impact. The study by Lublin et al. is smaller (n=224) and analyses participants from placebo arms of clinical trials in the 1980s-90s. The participants were, on average, older than the natural history cohort (mean 35.2 years) and included both relapsing and progressive phenotypes but had a longer (retrospective) follow-up (mean disease duration <15 years). Assessing EDSS before and after a relapse, Lublin et al. found that 42.4% of participants in the placebo arms of clinical trials had an increase of ≥0.5 EDSS points, with 28.1% increasing by more than 1 EDSS point. Conversely, however, this means that the majority had no change (38.4%) or improvement (19%) in EDSS following relapse. In addition, the post-relapse EDSS was calculated an average of 2 months, and a maximum of 4 months, following the relapse. This is concluded as irreversible disability for those who had an increased EDSS score, yet it has since been shown that such short-term evaluations are unreliable[35]. Indeed, even prior to the Lublin study, it had been demonstrated that half of those patients who experienced 3- or 6-month disability 'progression' had reverted to a non-progressed state by the end of a 2-year clinical trial[36]. Kalincik et al. take this further by analysing data from 16636 patients from the global MS Base database who had reliable examination follow-up over an average of almost 6 years[35]. 'Reversal' of EDSS increases was common and dependent on the time to confirmation, with the longest duration being the most reliable i.e. likely to persist after 5 years. Almost a third of patients with increased EDSS scores, following relapse, 'confirmed' at 3 months had reverted to their baseline by 5 years; for those confirmed at 2 years, only 11% reverted. This, therefore, calls into question the reliability of the Lublin et al. conclusions regarding the impact of relapses. Similarly, residual deficits or changes in deficit are measured and reported in EDSS units less than 0.5 e.g. 36.5% of patients had

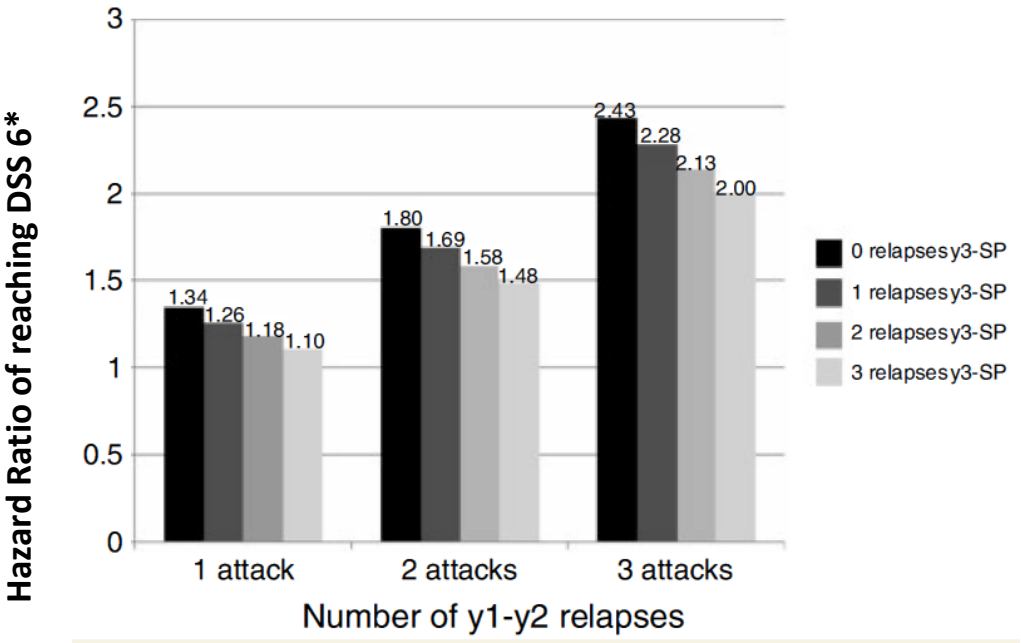
a mean residual deficit of 0.24 EDSS points, with an average worsening of 0.24-0.57 EDSS points per relapse. The EDSS scale is non-linear and categorical, resulting in an uncertain interpretation of these values which have no clinical correlate.

As well as the (lack of) effect of relapses on progressive MS disease courses, the Ontario group and others have also analysed their relationship with disability in patients with a relapsing-remitting course. Certainly, relapse frequency within the first 2 years appears to correlate with disability milestones. Those who had 1 relapse in the first 2 years reached DSS 6, 8 and 10 at 7.6, 12.8 and 20.3 years later, respectively, in comparison to those who had more than 3 relapses in the first 2 years after onset[32]. Similarly, relapses within the first 2 years were associated with an increased risk of developing SPMS (HR 1.1, p=0.003), with relapses in year 2 more predictive than those occurring in the first year following disease onset. However, the frequency of relapses after year 2 had no predictive value with regard to disability milestones or the time to secondary progression (SPMS). Relapse frequency diminishes over time, tending to be highest during years 1-2. Perhaps, then, it is the rate of relapses, rather than the total number, which is of most relevance. Additionally, the neuroanatomical location of the initial relapse is relevant, with brainstem events associated with a shorter time to DSS 6 and 8.

As anticipated, there is a directly proportional relationship between the number of relapses in years 1-2 and time to onset of SPMS. Similarly, the longer the duration between the first and second relapses, the longer the time until SPMS develops. Most interestingly, however, relapses occurring from year 3 until the onset of SPMS had an inverse relationship with the time to onset of SPMS. Those having no relapses between year 3 and the onset of SPMS developed SPMS at a mean of 8.2 years from disease onset whilst this occurred after a mean of 13.6 years in those have ≥3 relapses after year 3 ($p<0.001$)[32].That is, relapses after year 3 appear *protective* against the onset of SPMS, as shown in **Figure 0-2**. The Hazard Ratio of developing SPMS in those having 5 relapses after year 3 (0.45) was nearly half that of those having a single relapse in the same period (0.85). Similarly, the risk of reaching DSS 6 *reduced* with an increased number of relapses between year 3 and the onset of SPMS. LeRay et al. found differently, however: among patients still relapsing despite reaching DSS 3 (i.e. not

developed SPMS), the time from DSS 3 to DSS 6 was longer (12 years) in those without relapses after DSS 3 than those who had relapses (9 years) but the difference was not statistically significant (p=0.677)[17]. However, whether relapses occurred or not in those with RRMS after reaching DSS 3, the time to DSS 6 was significantly longer (9 years) than those with established SPMS who had reached DSS 3 (6 years) suggesting that conversion to SPMS was a more important influence on disability progression than the occurrence of relapses.

Figure 0-2: Relapses and risk of disability over time (From Scalfari et al.[32])



*in comparison to 0 relapses

This apparent dissociation between relapses and disability progression has been replicated in other studies, most eloquently in the British Columbia natural history cohort[33,37]. Firstly demonstrating that relapse rates naturally decline with time and that older patients have fewer relapses[37], Tremlett et al. also subsequently showed that relapses in the first 5 years have significantly greater impact on the risk of reaching EDSS 6 and the time to SPMS in comparison to relapses which occur >10 years after onset[33]. Both studies included a cohort of 2477 after exclusions, with between 16.8[37] – 18%[33] having been treated with a DMT. Notably, DMTs were initiated a mean of 17.8 years after onset of MS (range

2-49 years) and the proportion of time on treatment was small (<3% of follow-up) and analyses excluding these patients did not alter the results. **Figure 0-3** demonstrates the effect of time and duration of disease on relapse rates.

Figure 0-3: Relapse rates A) with disease duration B) with age (from Tremlett et al.[37])

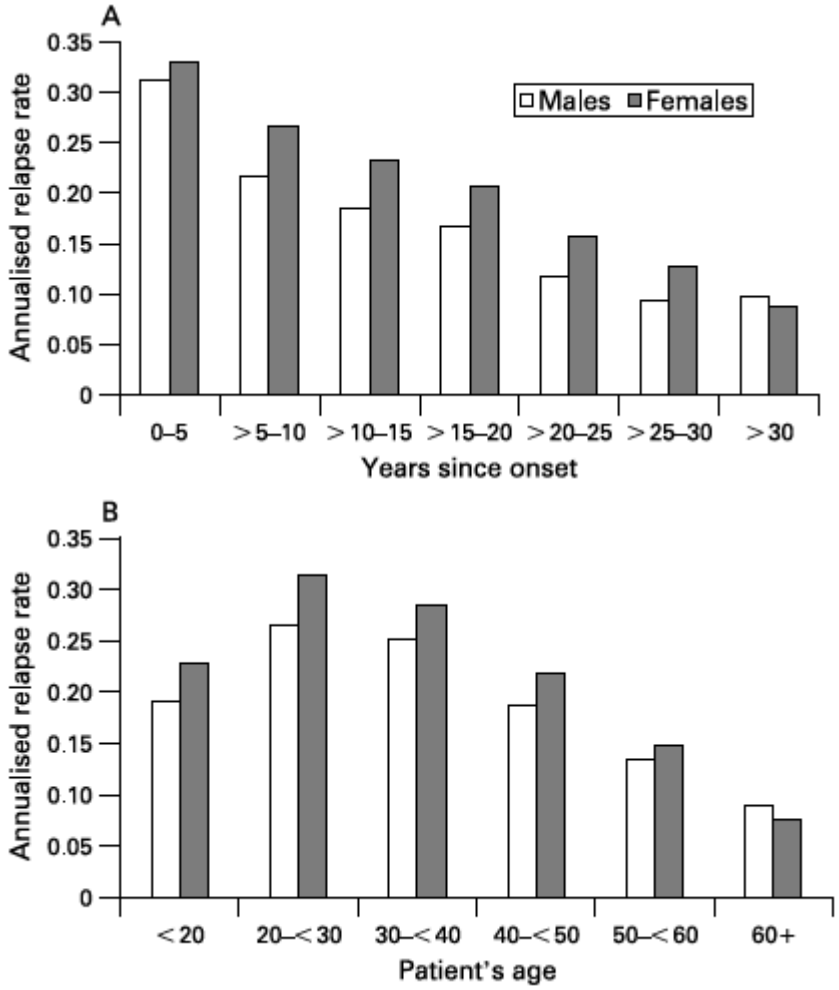


Figure 0-3 demonstrates the decline in relapse frequency both with increasing duration from disease onset and the age of the patient at the time. This includes 20.6 years of follow-up and >11500 relapse events. The mean time from onset to first clinic visit was 12.1 years and the concern over retrospective labelling of relapses is raised by the authors, meaning potential underestimation of relapse frequency. This bias is, arguably, consistent throughout all epochs studied hence should be balanced across the cohort however. Additionally, analysis of the subgroup seen within 5 years from onset (n=626, 25%) showed similar results,

albeit with increased relapse frequency. Overall, the relapse rate reduced by 17.5% every 5 years, increasing with increased patent age. Additionally, 77.3% of those with more than 5 years follow-up in the relapsing-remitting phase (n=2189) experienced a 5-year relapse-free period during the relapsing-remitting phase[37]. This is in the absence of any disease-modifying treatments for the vast majority.

This natural decline in relapse rate and notable prolonged relapse-free periods must be considered when interpreting the impact of DMTs in both research and clinical practice, particularly where there is no defined control arm. That said, the patients who had prolonged relapse-free periods took more than double the time (23 years), on average, to develop SPMS in comparison to those without these periods (10.9 years to reach SPMS) [$p<0.0005$]. This is perhaps evidence toward relapses contributing to the progressive phase or simply that disease severity is variable within a population and those with fewer relapses have a milder form of the disease. Patients under 20 years of age had a small increase in relapse frequency over time, peaking between ages 20-30: the decline otherwise continued thereafter. The rate of decline in relapses differed depending on the age at onset of MS. For every 5-year epoch, relapse rates decline by 30.5% for those with onset beyond 40 years of age, 22.9% for those aged 30-40, 17% for 20-30 year-olds and 7% for those with onset before age 20.

Along with the declining frequency of relapses with time, Tremlett et al. also found evidence that relapses occurring later in the disease course have a gradually diminishing contribution to the risk of disability or entering the progressive phase (SPMS). A relapse within 5 years of disease onset was associated with an increased hazard in early disease progression, within that 5 years, of 48% (95% CIs 37-60) for requiring a walking aid (EDSS 6) and 29% (95% CIS 20-38) for developing SPMS in comparison to the situation where no relapse occurred in this time. In contrast, a relapse occurring after 10 years post-onset increased the risk by just 12% (8-17) and 8% (4-11) for these endpoints respectively in the long-term if they had not already been reached[33] as shown in **Table 0-2**. Disease onset refers to the first likely symptoms rather than the date of diagnosis, notably.

Table 0-2: Hazard of reaching endpoints if relapse occurred in time-frame in comparison to no relapse occurring from disease onset (Adapted from Tremlett et al.[33])

	Short-term Hazard (0-5yrs) of		Medium Term Hazard (5-10yrs) of		Long-term Hazard (>10yrs) of	
	EDSS 6	SPMS	EDSS 6	SPMS	EDSS 6	SPMS
Relapse at 0-5 yrs	1.48 (1.37-1.6)	1.29 (1.2-1.38)	1.25 (1.17-1.34)	1.11 (1.06-1.17)	1.10 (1.04-1.16)	1.02 (0.98-1.07)
Relapse at 5-10 yrs			1.31 (1.22-1.42)	1.23 (1.15-1.31)	1.07 (1.01-1.13)	1.06 (1.02-1.11)
Relapse after >10yrs					1.12 (1.08-1.17)	1.08 (1.04-1.11)

In this analysis, if EDSS 6 was not reached within 5 years after onset, every relapse within the first 5 years was associated with a 25% (95% CIs 17-34) increased hazard of this occurring in the next 5 years. Similarly, the risk of developing SPMS was increased by 23% for every relapse during this time in comparison to the scenario where no relapse occurred in the first 5 years (p=0.05). The risks dropped to a 12% increase (95% CIs 8-17) for EDSS 6 and 8% (95% CIs 4-11) for SPMS if the relapse occurred after 10 years. There was no significant difference in the hazard of reaching EDSS 6 or SPMS in the long-term (>10 years post-onset) whether a relapse occurred earlier or later in the disease course. Increased frequency of relapses within the first 5 years was highly associated with an increased risk of reaching EDSS 6 and SPMS 10 and even 30 years later, however, including after adjustment for gender, age at onset and symptoms ($p<0.0005$ for both EDSS 6 and SPMS using the log-rank test).

The association of relapses with disability and SPMS development was age-dependent, being greatest (i.e. higher association with disability endpoints) in patients aged less than 25 at disease onset and least in those older than 35. In

those over 35 years of age, there was no statistically significant increased risk of reaching EDSS 6 if a relapse occurred more than 5 years after disease onset or developing SPMS if it occurred 10 years after disease onset. Relapses occurring within the first 10 years of disease of those with onset at less than 25 years of age have a significantly greater impact on their risk of disability than in a patient whose disease started after 35 years of age. Relapses occurring during the secondary progressive phase had little or no impact on the future risk of reaching EDSS 6 if this had not already occurred by the time of SPMS onset. Overall, using a multivariate model, women were at a 30% (95%CIs 16-42) and 31% (23-39) lower hazard of reaching EDSS 6 and SPMS respectively in comparison to men. There was a 17% (12-23) increased risk of reaching EDSS 6 and 25% (21-28) for SPMS for every 5-year increase in age at disease onset.

For all of these results, it is notable that a substantial proportion of the cohort had not reached the defined endpoint by study end, meaning they were censored from the analyses: over 70% had not reached EDSS 6 and almost 50% had not developed SPMS. Much of the findings are therefore based on a minority of patients who reached the proposed endpoints and therefore had a relatively severe disease course. The analyses also suffer from attrition of the at-risk cohort with time, such that, for some analyses e.g. endpoints beyond 20 years, the results are based on a cohort size of double figures only. Of course, the true impact of relapses on the life of a patient could only be assessed if the entire cohort were followed until all reached the pre-defined endpoints, but this is an impractical goal and these data provide the best longitudinal analyses available.

In summary, these findings suggest that a relapse within the first 5 years from disease onset confers an increased risk of developing irreversible disability or SPMS in the short-term, but the risk of an early relapse has much less impact on long-term future risk. A higher frequency of relapses within the first 5 years is associated with a higher risk of disability in the long-term, however. For those with disease onset before age 25, the risk of irreversible disability and SPMS remains increased even from a relapse occurring 10 years after diagnosis but the magnitude of this risk declines with time. In contrast, those with disease onset

after 35 years of age have less increased risk of irreversible disability or SPMS from a relapse occurring after 5 years from disease onset and no statistically significantly increased risk for a relapse occurring after 10 years from onset. Relapses during the secondary progressive phase do not appear to have a discernible impact on disability in general, as measured by EDSS at least. These data would suggest that targeting relapse-reducing therapies at younger patients in the earliest stages of their disease is likely to have most impact, in the short-term at least, whilst the benefits of aggressive treatment in older patients with long disease duration are much less certain and relapse-reduction in SPMS is unlikely to outweigh any risk of treatment, if maintenance of mobility or preventing increased EDSS are the goals. The impact of relapses, in terms of disability or risk of SPMS, being highest in the immediate aftermath of a relapse (within 5 years of the event) suggests that the British Columbia studies may be capturing relapse-related disability (albeit irreversible) rather than neurodegenerative progression where, as above, a causative relationship with relapse remains unproven.

MS treatment and trial considerations

In the early 19th century, multiple sclerosis was recognised as one of the 'paraplegias', a sub-class of the known 'nervous disorders' at the time. Apoplexy, epilepsy, neurosyphilis, congenital idiocy and brain fever were the other classifications from which the paraplegias were differentiated by the occurrence of progressive weakness. Whilst Charcot is credited with the first recognition of MS, his attempts at treatment were unsuccessful[38]. Meat diets, bleeding, cooling, strychnine, quinine, belladonna, arsenic, atropine, ergot and alkaloids were all cited as ineffective by William Moxom in 1875, who had first described MS in the UK. The evidence on which this conclusion was based, however, is unclear. In Germany, electrical and magnetic stimulation were used with, presumably, equally unsatisfying results given the subsequent departure from this. The first critical review of MS treatments was undertaken by Russell Brain in 1930, where he describes the understanding of the condition at the time and the past experience of treatments used[39]. As well as some of the treatments considered by Moxom previously, vaccines and other pyrogenic agents are included but similarly considered ineffective. An interesting conclusion of

Brain's, which still has some relevance today perhaps, is that 'the multiplication of remedies is eloquent of the inefficacy'. It is certainly the case that treatments today are more effective but there remains no panacea.

Whilst antibiotics and antifungals were also trialled, ineffectively, after their widespread introduction in the 1940s, no major advances were made until the introduction of randomised control trials in the 1960s. The first RCT in MS assessed the utility of aspirin and steroids in a 3-armed placebo-controlled study[40]. In fact, this found no difference in outcomes between the three arms, but was primarily testing the hypothesis that MS was an allergic, inflammatory condition. Brown et al. published detailed and thoughtful guidelines on the design and conduct of clinical trials in MS in response to the perceived inadequacy of previous research[41]. This international panel recognised the difficulties in MS clinical research highlighted previously by Schumacher[42] and, indeed, some of these remain today. The lack of diagnostic precision, unpredictable course and difficulties in quantifying disease severity and maintaining large groups of patients on standard treatments for extended periods were relevant both then and now. Specifically, the inability to prove the diagnosis in life was raised by Brown et al., who were concerned that poorly designed studies were wasteful and lead to false hopes for patients and clinicians. Importantly, they considered the ultimate goal of MS treatment to be the prevention of its first clinical manifestation or the complete prevention of even the subclinical form. With a pragmatic approach, however, they recognised the lack of understanding of the aetiology of MS and hence considered treatment and prevention of relapses and progression as more achievable goals. This has really set the tone for future research, particularly the pivotal DMT trials. Randomisation, blinding, sample size, inclusion and exclusion criteria, follow-up of dropouts, statistical analyses, funding and collaboration are all discussed as issues to be considered in the planning of MS trials in this detailed appraisal. Interestingly, the paper also suggests the need for publication of all data related to a trial (as a supplement); this anticipates the very recent advocates of full disclosure but, as yet, is not widespread.

Despite the detailed guidance from Brown et al in the 1970s, modern clinical trials of MS treatments have been criticised on a number of fronts, both in their design and analyses[43,44]. The natural inter- and intra-patient variability in the

course of MS results in a higher risk of analysis errors. The falling rates of relapses in modern cohorts in both treatment and placebo arms, for a variety of reasons as described below, necessitates larger cohort sizes to demonstrate statistically and clinically significant differences. This requires recruitment from multiple specialist centres, often internationally, and hence added complexity in the design and conduct of the trial as well as increased heterogeneity in both the study participants and healthcare teams. The sample size must be even larger when the comparator is a known effective treatment, rather than placebo, in order to demonstrate differences against a new treatment under study. Whether the trial is placebo-controlled or against an active comparator will also affect the population included: those with milder disease are more likely to consider a placebo-controlled trial whereas those with active disease, and their physicians, will be keen to ensure they are definitely receiving a treatment. Similarly, in modern trials, the availability of effective treatments results in a larger proportion of non-responders recruited to trials with active comparators only when they have already failed on a licensed therapy. This may bias the cohort towards poorer outcomes. Additionally, the cost of such huge studies cannot be borne by governments or academic institutions, resulting in a frontline role for the pharmaceutical industry. The commercial trial industry is now the main vehicle for large-scale international research trials but leaves any analyses and conclusions open to criticism given the inherent bias when industry has any role other than purely funding.

A number of potential errors in the design and analysis of clinical trials in MS have been highlighted and these must be borne in mind when appraising the evidence of DMTs currently in use or purported for future use. Strictly speaking, trials provide no evidence for the effectiveness of treatments in individuals who have characteristics other than those included in the study[45] but, in practice, the results are generally extrapolated to the general MS population. This is, arguably, invalid but even the purported effect sizes in the target population will not necessarily be reflected in a real-world population: Montalban differentiates this as the difference between *efficacy* and *effectiveness*[44]. Efficacy is the effect demonstrated in a clinical trial, whilst effectiveness corresponds to the treatment effect in everyday practice. Of course, in everyday practice, the population, intervention, monitoring and quality of care are not

standardised in the same way as a research study. Misdiagnosis, prescription errors, concordance and failure of follow-up are just some of the issues which can arise to jeopardise the replication of trial results in the real world. To some extent, this is addressed by intention to treat analyses, rather than per protocol, but even this does not account for the daily shortcomings of clinical practice and less motivated patients in real life. Whilst a (theoretical) treatment may be *efficacious* in the vast majority of the trial population who adhere to the regime, even small increases in the proportion with misdiagnosis, prescription errors, concordance issues and inadequate follow-up will have an additive effect that soon reduces the *effectiveness* of the treatment to well below that found in the trial. In addition, the patient populations may differ in international trials, particularly where treatment is only available to those who participate in the trial or have to pay for it otherwise.

The statistical analyses and interpretation of results from clinical trials has become increasingly complex and more challenging for the non-statistician clinician to translate into meaningful conclusions for both themselves and their patients. The p-value is a case in point. The statistical significance of a result relates to its believability – it is the *effect size* which relates to its clinical importance. Trials with large effects of marginal significance or significant effects of marginal importance should both be judged as providing equivocal evidence[43]. The utility of the p-value is in setting a pre-experiment significance value rather than the significance of a result once the data has been collected. The p-value has gained increasing importance in published results but its origins were not clearly intent on this use. A p-value <0.05 has been widely agreed as an acceptable level of certainty of excluding chance findings in medical research and can be interpreted as meaning there is a less than 0.05 probability (5% or 1 in 20) that the value found (or a value more extreme) occurred by chance. Crucially, a p value of 0.05 does not, however, imply that there is a 95% chance that the effect is real.

Over time, it seems the p-value and the alpha error rate (Type 1 error) have become considered interchangeable but they have different origins and are conceptually different. The alpha (Type 1) error rate is the probability of incorrectly rejecting the null hypothesis i.e. concluding there is a difference between two groups when, in fact, there is not. An observation with a p-value

equal to 0.05 actually has an alpha error of 13% in a one-sided hypothesis test and 21% in a two-sided test. The p-value represents an area under a curve whilst an observed value occurs at a single point and, hence, Goodin advises that a p-value of <0.01 for an observed value is actually required to ensure the alpha error rate is less than the conventional 5% and p-values between 0.01 and 0.05 should be considered marginally significant[45]. The alpha error rate and its beta counterpart (the probability of incorrectly accepting the null hypothesis i.e. concluding there is no difference when there actually is) were devised as part of the Neyman-Pearson hypothesis test in 1928[46]. The alpha and beta rates were created to define 'critical regions' in the distribution of any summary statistic (e.g. mean value), within which any value was prone to either of these errors. Indeed, if a value fell into these critical regions then the null hypothesis was to be rejected and the alternative hypothesis accepted; equally, values outwith the critical regions suggested the null hypothesis was correct and that the alternative hypothesis should be rejected. On completion of the experiment, (Fisher's) p-value was used to report whether the result fell within the critical region or not; it was not intended to show where in the critical region it fell, however. Neyman stated that there is no objective difference in a p-value of 0.04 in comparison to a p-value of 0.00001 as both are significant at alpha equal to 0.05. Indeed, the 'equal to' is contentious also. A p-value can be reported precisely (p=x) or imprecisely (p<x). The imprecise value conveys meaning to the value and anything more extreme, whilst the precise value only represents the border of the tail on a distribution curve – it does not imply anything about more extreme values. Using the precise value overestimates the case against the null hypothesis (i.e. suggests there is a difference between the groups). Goodman exemplifies this with clinical results where p is given as < 0.05 having a less than 3% chance of the null hypothesis being true, as opposed to 25% where a p value is given as equal to 0.05. Again, this difference diminishes when p values ≤0.001 are used but values between 0.001 and 0.05 may lead to overestimation of the plausibility of results[46]. The beta error rate relates to sample size and power and can be considered in trial design, with planned recruitment and stated effect sizes considered relevant, but cannot be adjusted for after the collection of data other than stating the study may have been underpowered.

There are other statistical pitfalls to be wary of which are relevant to all medical research but particularly so to MS studies and have occurred in published literature. Post-hoc analysis is fraught with statistical pitfalls and, in general, should be interpreted with extreme caution. This is where analyses are considered only after the data has been collected and is prone to 'fishing' for positive results or correlations, particularly if the primary endpoint was not met. If an analysis has not been planned *a priori* then the cohort and data are unlikely to have been collected in a valid way to answer the question which arises later. Subgroup post-hoc analysis can detect statistically significant differences in groups separated on entirely arbitrary grounds[43]. Additionally, when more than one hypothesis test is undertaken on the same dataset e.g. multiple comparisons to evaluate for significant differences, this must be considered in the resultant p-value. If multiple comparisons are made, the chance of rare events occurring increases and hence so does the risk of making a Type 1 error i.e. incorrectly rejecting the null hypothesis. The Bonferroni adjustment can account for this to some extent and requires that the desired alpha rate is divided by the number of tests conducted. Thus, if 0.05 is taken as the acceptable significance level (for a single result being tested for significance), this must be reduced if more than one hypothesis is being tested in the same data. If five hypotheses are being tested, for example, only an alpha rate of (0.05/5) 0.01 should be considered statistically significant.

Regression to the mean is another statistical pitfall to which MS studies are potentially prone. Particularly in earlier studies, cohorts were specifically selected for their high disease activity relative to the general MS population. This aimed to ensure adequate disease activity would occur over the trial period in order to attain pre-set endpoints: an active group will have more activity and hence require shorter follow-up to demonstrate differences between groups. Indeed, this holds true for placebo-controlled trials but is invalid in any study without a control group. By definition, the individuals selected for these studies are unusual in their phenotype but, as a universal truth, they will tend to become more 'usual' over time i.e. 'regress to the mean'. This must be considered in any single-arm or crossover study as, without any intervention, those with initially high disease activity will tend to have less over time. Similarly, this means one cannot compare intra-group changes between two

groups. It is not valid to compare the change within a group with that of another group undergoing a different intervention – only the difference between the groups can be considered relevant. Goodin cites the European North American Comparative Efficacy (EVIDENCE) trial to demonstrate this error[43,47]. Participants in this study, comparing Avonex® with Rebif®, were offered open-label Rebif® at the end of the trial after it was demonstrated this was the more effective of the two. Relapse rates in the last 6 months of the trial were compared with those in the open-label phase and demonstrated a significant ($p<0.001$) reduction in relapses in those who switched from Avonex® to Rebif® but a smaller reduction in those remaining on Rebif® continuously ($p=0.03$). However, the comparison was between the reduction in relapse rate in each group, rather than comparing the reduction between groups and, in fact, there was only a minimal difference between the groups when analysed in this way ($p=0.05$). Only this difference is a valid comparison and the previously stated p-values are irrelevant and misleading.

Studies where inclusion is dependent on the occurrence of an intervention at a time after recruitment, in comparison to no intervention, are at risk of immortal time bias (also known as survivor treatment selection bias). In this error, the participants in the intervention arm will, by definition, have the intervention whilst participants in the non-intervention arm have no requirement for this. Therefore, any event occurring between recruitment and the intervention which prevents those in the intervention arm from having it (e.g. death) will bias the study toward a milder intervention group - death would not exclude those in the non-intervention group from the analysis. Participants in the intervention group who die between recruitment and the intervention will not be included in the analysis if the occurrence of the intervention is necessary for inclusion. The intervention group, when considered in this way, are therefore 'immortal' for the period between recruitment and the intervention, whilst those dying in the non-intervention group will be included in the analysis and hence provide the appearance of a more severely affected cohort. This holds true for any event of interest being recorded between recruitment and the intervention, if the event may influence the likelihood of a participant receiving the intervention. This error can be eliminated by setting the time of intervention as the beginning of the study or by using time-dependent covariates in any survival analysis[48].

The lack of multi-arm comparator studies of DMTs in MS necessitates extrapolation of the comparative efficacy of each treatment, but this is fraught with potential problems. Ideally, all available DMTs (or a group of them) would be subject to a multi-arm randomised trial in order to demonstrate their comparative safety and efficacy. The different methods of administration would make blinding almost impossible, however, as exemplified by the CARE-MS trials comparing Alemtuzumab, an annual intravenous infusion, with Rebif, a thrice-weekly subcutaneous injection[49]. Similarly, the differing risk profiles would make recruitment challenging given the potential of randomisation to treatments with significantly more safety issues or perceived reduced efficacy. Sample sizes would have to be well in excess of the usual two-arm trials, which are already necessarily large and complex, particularly with the falling levels of disease activity seen in recent trials[44]. In addition, this would require significant financial backing and the agreement of all manufacturers of the drugs included, in the knowledge that their product is at risk of obsolescence if definitively proven less safe and/or effective relative to the other options included. It seems unlikely that such a study will be undertaken given these design and practical barriers. Much debate is undertaken, therefore, on how best to decide which is the safest and most effective treatment for the patient in clinic. Of course, even such a multi-arm theoretical trial would not answer this for an individual patient given the reliance on group effects to demonstrate safety and efficacy and the lack of individualisation of responses. The role of individualisation of prognosis and treatment is a new frontier in medicine and studies in MS are already trying to address this.

Whether inter-study comparisons of DMT trials in MS are valid, or at least reasonable, remains a matter of debate[50,51]. Some suggest that the relatively standardised inclusion criteria and measures of disability and disease means patients included in many modern MS treatment trials can be compared[50]. Similarly, a fairly narrow range of outcome measures have been used in the pivotal DMT trials (e.g. EDSS, ARR, MRI) and, hence could be considered comparable also. Post-hoc approaches to allow comparative analysis of independent DMT trials i.e. network meta-analysis have also been used and are considered a valid tool for this purpose[50,52,53]. From a practical point of view, the

costs of the studies and development of new therapeutics is unlikely to be borne by anyone other than multinational pharmaceutical industries and the risk of their disengagement from this, in the event of a move toward multi-arm head-to-head trials, could jeopardise future research endeavours. Finally, it is recognised that trial populations do not represent fully the target treatment population in the 'real world' and, hence, large post-licensing registries with appropriate follow-up data and pharmacovigilance would provide much of the evidence required to compare DMTs in terms of efficacy and safety. Indeed, if these data were included as part of a multi-centre (or multi-national) prospective observational study the results would be potentially comparable to that of a randomised trial[54].

Others believe efficacy comparisons using independent DMT trials are not valid[51]. Arguably, the change in the natural history of the condition, with the broadening inclusion of milder phenotypes, precludes rational comparisons between trial cohorts in the early 1990s with those today for example. Similarly, as well as the definition of the disease changing, the definition of disease activity is variable between independent trials. Newer trials have more emphasis on imaging (and with more powerful scanners with more available sequences) than in the past; relapses are defined differently in each trial, with the most recent being more stringently-defined in general. Stricter definitions of what constitutes a relapse e.g. whether an examination or EDSS change is required, will lead to fewer recorded relapses than in studies without this. Additionally, the frequency of study visits is proportional to the number of relapses documented, given the closer monitoring and questioning of participants. Measurements of disease progression are inadequate and can be misleading over short periods, yet 3- or 6-month disability (EDSS) worsening is the mainstay to define this in large DMT trials, despite there being a known reversibility of 'progression' when defined in this way[35]. When MRI outcomes are considered, there is wide potential variability in their acquisition, analysis and interpretation depending on the methods used between studies, making direct comparisons often invalid. This is particularly the case for MRI measurements of atrophy.

Indeed, combined outcome measures are now suggested in order to capture all known disease activity. The concept of 'No Evidence of Disease Activity' (NEDA) assumes we know and can measure all MS-related sequelae but generally

describes no evidence of relapses, progression or MRI changes (NEDA-3); this can be extended to include MRI-estimated brain atrophy (NEDA-4) and neurofilament levels (NEDA-5) and could likely be extended further to include cognition and the myriad of other factors considered relevant. Similarly, No Evidence of Progression or Active Disease (NEPAD) extends to include upper limb function (9HPT) and Progression Independent of Relapse Activity (PIRA) aims to identify non-inflammatory disease worsening thereby potentially identifying the true neurodegenerative aspect.

However, in lieu of either a multi-arm randomised trial of DMTs or a well-established observational registry with long-term follow-up and sample sizes to make conclusive comparisons, patients and clinicians can only rely on the evidence available and must make pragmatic decisions based on this. The pivotal DMT trials and attempts to extrapolate valid comparisons between them are considered in detail below.

DMT Trials

In the era of evidence-based medicine, the gold standard for disease intervention is demonstration of efficacy in an agreed outcome measure against either placebo or an active comparison treatment in a randomised, double-blind controlled trial. **Tables 0-3, 0-4** and **0-5** outline the pivotal trials for currently licensed RRMS DMTs and **Table 0-6** lists emerging therapies which have been studied but not necessarily licensed for use yet. Most of these qualify as Class I evidence for the efficacy of the intervention being studied, save for the Alemtuzumab and Ocrelizumab trials where blinding was not realistically practicable or ethically defensible with the availability of proven therapies. That said, these data summarise the best available evidence for DMTs in RRMS. It is widely agreed that cross-trial comparisons are not valid, and the variation in numbers, cohort characteristics, definitions and outcome measures described here reinforce this point, as discussed previously. Additionally, the trials outlined have been undertaken between the early 1990s and late 2010s i.e. over a 20-year period during which a multitude of other societal and medical changes have occurred. Indeed, the diagnosis of MS has notably changed over time, with each iteration of diagnostic criteria generally allowing inclusion of patients with earlier (milder) disease than previously. There have been comparator trials, as outlined in the subsequent table, allowing clear statements about the superiority of some DMTs over others, but most DMTs have not been compared in head-to-head trials.

Inevitably placebo-controlled trials are compared in practice, in the absence of other evidence. The raw outcome comparisons, comparing *relative* ARR reduction as the typical primary outcome measure, are usually used and quoted, but some argue that this inflates apparent efficacy outcomes and absolute differences should be used[55]. Relative ARR risk reduction figures suggest natalizumab as the most effective DMT (68% reduction) versus placebo followed by cladribine (57.6%). DMF and Fingolimod had similar relapse rate reductions, as did Teriflunomide and Plegridy®. The other interferons and Copaxone® have relapse reduction rates of around 30%, other than Avonex® for which this was only demonstrated in the per-protocol analysis. Clearly, as stated, this is an artificial and simplistic view, and relapse reduction is not the only measure of disease activity and, arguably, not the most relevant to patients or society.

Additionally, tolerability and safety are important issues for patients and hence real-world concordance. From these pivotal trials, Copaxone® has the best tolerability profile (1.6% discontinued) and the oral treatments the worst.

The comparator trials (**Table 0-7**) provide the best evidence as to the efficacy of one DMT versus the other but, again, cross-trial comparisons are not valid. From these we can however conclude that:

- Alemtuzumab is more effective than Rebif® (Further details in **Table 0-5**)
- Rebif® is more effective than Avonex®
- Betaferon® is more effective than Avonex®
- Fingolimod is more effective than Avonex®
- Ocrelizumab is more effective than Rebif® (Further details in **Table 0-6**)
- Copaxone®, Betaferon®, Rebif®, Teriflunomide and DMF are of comparable efficacy
- Natalizumab in addition to Avonex® is more effective than Avonex® alone
- Natalizumab in addition to Copaxone® is not more effective than Copaxone® alone

There are clearly caveats to the above which is an over-simplification, and there are some obvious potential contradictions with the placebo-controlled trial outcomes e.g. the addition of Natalizumab to Copaxone not improving efficacy. There may be reasons for this which are not yet understood, in addition to the small study sample size, and these data, despite their problems, represent the gold standard available.

Interestingly, despite the caveats discussed, meta-analysis of all available trial data seem to bear out the 'league' of efficacy suggested by crude comparisons of the pivotal trials[56]. From this analysis, Alemtuzumab and Natalizumab appear most effective in most measures, followed by Fingolimod and DMF, and then the injectable therapies. Teriflunomide is of mid-range efficacy in terms of 3-month disability progression, but similar to the injectables in terms of relapse activity. Ocrelizumab and cladribine were not included in this analysis.

In summary, a number of factors function in patients' treatment choices but the available data provides clinicians with some guidance on the most appropriate treatment advice for an individual. Both efficacy and safety/tolerability are

relevant in deciding on which DMT to use, with the frequency of discontinuation very relevant given the impossibility of effectiveness if patients choose not to use the treatment. **Table 0-8** outlines a summary of efficacy and discontinuation rates from the pivotal trials and suggests favourability on this measure for the most efficacious treatments, in terms of relapse reduction, as they are also well tolerated, often as a result of the short courses of treatment required (in the case of Cladribine and Alemtuzumab for example).

Table 0-3: Pivotal DMT trials - First-line Injectable treatments

	Avonex®		**Betaferon®**		**Rebif®**		**Plegridy®**		**Copaxone®**	
DMT/Comparator	Avonex®	Placebo	Betaferon® (8MIUs)	Placebo	Rebif44	Placebo	125µg 2weekly	Placebo	Copaxone 20mg/day	Placebo
Trial (Year)	1996 (MSCRG)[57]		1993[58]		1998 (PRISMS)[59]		2014 (ADVANCE)[60]		1995 (MSSG)[61]	
Trial details										
N [analysed]	85	87	115 [58]	112 [56]	184	187	438	456	125	126
Primary outcome	Time to ↑EDSS≥1		ARR / Proportion relapse-free MRI: %Δlesion area		Relapse count		ARR		Relapse rate	
Duration	104 weeks		2yrs [5yrs Follow-up]		2yrs		48 weeks		2yrs	
Inclusion Criteria	CDMS>1yr EDSS 1-3.5 18-55yrs ≥2 relapses in 3yrs		CDMS or 'lab supported'>1yr EDSS 0-5.5 18-50yrs ≥2 relapses in 2yrs		CDMS or 'lab supported'>1yr EDSS 0-5 ≥2 relapses in past 1yr		McDonald criteria (2005) EDSS ≤5 18-65yrs ≥2 relapses in 3yrs, at least 1 in last yr		EDSS 0-5 18-45yrs ≥2 relapses in 2yrs	
Exclusion Criteria	'Any disease compromising organ function' Pregnancy or 'unwillingness to practice contraception' Prev. immunosuppression		AZA or Cyclophosphamide use previously		Any previous immunotherapy in past 1yr		Previous IFN use		No relapse/steroids 30 days before study Diabetes Pregnancy/lactation No Aspirin/NSAIDs during trial	
EDSS frequency	6-monthly		3-monthly		3-monthly		3-monthly		3-monthly	

	Avonex®		Betaferon®		Rebif®		Plegridy®		Copaxone®	
DMT/Comparator	Avonex®	Placebo	Betaferon® (8MIUs)	Placebo	Rebif44	Placebo	125µg 2weekly	Placebo	Copaxone 20mg/day	Placebo
MRI frequency	Annual		Annual (52 pts had 6-weekly for 2yrs)		Every 6/12 (205 pts had monthly scans for first 9/12)		6/12		N/A	
ITT Design?	Yes		Yes		Yes		Yes		?	
Relapse definition	≥48hrs and EDSS↑		≥24hrs and new sign		≥24hrs, stable for 30/7 before		≥24hrs, stable for 30/7 before and signs		≥48hrs and EDSS↑, stable for 30/7	
Progression definition	↑EDSS ≥1 @ 6/12		↑EDSS>1 @90days		↑EDSS>1 @3/12		EDSS ↑>1 if baseline EDSS >1 or ↑>1.5 if baseline EDSS 0		↑EDSS>1 @3/12	
Baseline Characteristics										
Mean age	36		35.5		35.6 (Median)	34.6 (Median)	36.9	36.3	34.6	34.3
% Female	72-75		68-72		66	75	71	72	70.4	76.2
%White	92-93		93-94		N/A	N/A	81	82	93.6	94.4
Mean ARR	1.2		Relapses in past 2yrs		Mean 3 relapses in previous 2 yrs: ARR =1.5		Mean relapses in past 1yr (3yrs)		Mean relapses in 2yrs	
			3.3	3.6			1.6 (2.6)		2.9 (1.8)	2.9 (1.8)
Mean EDSS	2.3		3.0	2.8	2.5	2.4	2.5	2.4	2.8	2.4

	Avonex®		Betaferon®		Rebif®		Plegridy®		Copaxone®	
DMT/Comparator	Avonex®	Placebo	Betaferon® (8MIUs)	Placebo	Rebif44	Placebo	125µg 2weekly	Placebo	Copaxone 20mg/day	Placebo
Disease duration (years)	6.5		4.7	3.9	6.4	5.3	6.9	6.3	7.3	6.6
MRI Brain T2 data	N/A		N/A		N/A		Mean No. T2		N/A	
							48.7	50.6		
MRI Brain Gd+ data	0 Gd+ lesions		N/A		N/A		% with Gd+ lesions		N/A	
	47%	46%					35	41		
	>4Gd+ lesions						Mean No. Gd+ lesions			
	20%	18%					1.2	1.6		
Efficacy Outcomes										
Relapses										
ARR (95%CIs)	0.61 (ITT)	0.9 (ITT)	0.84	1.27	Mean relapses/pt		0.256* (0.206-0.318)	0.397 (0.328-0.481)	0.59*	0.84
					1.73	2.56				
Time to 1st relapse	Not significant		295days*	153days	Prolonged by 5/12 vs. placebo		N/A		287 (NS)	198
ARR Reduction	32%* (completers) [18% whole group]		33.9%*		33%* (21-44)		35.5%		29.7%	
									33% in EDSS 0-2@baseline	
% relapse free	N/A		36*	18	32*	16	81*	71	33.6 (NS)	27
Severity	N/A		Mod/Severe relapses (mean)		Mod/Severe relapses (mean)		N/A			

	Avonex®		Betaferon®		Rebif®		Plegridy®		Copaxone®	
DMT/Comparator	Avonex®	Placebo	Betaferon® (8MIUs)	Placebo	Rebif44	Placebo	125µg 2weekly	Placebo	Copaxone 20mg/day	Placebo
			0.23*	0.45	0.62*	0.99			N/A	
Disability										
% progression-free	78.8	66.6	5yr data (n=114)		N/A		99.93*	90	78.4 (NS)	75.4
			65	54						
Change in EDSS	↑EDSS>2.5		No difference after 2 and 3 yrs		↑0.24*	↑0.46	N/A		% improved (EDSS↓≥1)	
									24.8*	15.2
	2.4%	11.4%			95% CIs include 0				% no change	
									54.4	56
									% worse (EDSS↑≥1)	
									20.8*	28.8
MRI										
Gd+	0 Gd+ lesions		N/A		N/A		Mean no. lesions		N/A	
	71%	57%					0.2*	1.4		
	>4Gd+ lesions									
	6%	11%								
T2	%Δ lesion vol.		%Δ lesion area yr 2/3		‘Burden of disease’		Mean no. new or enlarging			
	-13.2%	-6.5%	↓0.1% ↓6.2%	↑20% ↑17.1%	↓3.8%*	↑10.9%	3.6*	10.9		
			6/52 MRI (n=52)		‘Active lesions’		Lesion volume Δ			
			Median new/yr		↓by 78%* vs placebo		↓0.26%*	↑0.77%		
			0.5	2						

	Avonex®		Betaferon®		Rebif®		Plegridy®		Copaxone®	
DMT/Comparator	Avonex®	Placebo	Betaferon® (8MIUs)	Placebo	Rebif44	Placebo	125µg 2weekly	Placebo	Copaxone 20mg/day	Placebo
T1	N/A		N/A		N/A		Mean no. lesions 1.8*	3.8		
Atrophy	N/A		N/A		N/A		N/A			
Safety Outcomes										
% Discontinued Rx	4	1	21 Withdrew due to SEs 8.7%	21 0.8%	Withdrew due to AEs 4.9%	 1%	4.7	1	1.6	?
Common AEs with Rx	Flu Sx (61%) Myalgia (34%) Fever (34%)		Fever (58%) Myalgia (41%) ↑LFTs (3pts withdrew) Injection pain Flu Sx esp. small pts		Injection site reactions Depression (2.7%) ↑LFTs (1%)		Injection erythema (62%) Flu Sx (47%) Pyrexia (45%) Headache (44%)		Injection site reaction (90%) Transient systemic reaction (15.2%) – can occur late Generalised lymphadenopathy (0.8%)	
% SAE	No difference		Nil		N/A		5% both groups		Nil	
Malignancy					Colon x1(?group)		Nil			
Infection					N/A		No difference			
Death	1 IFN pt - PTE		1 suicide ?group		1 suicide(placebo)		1	2		
Autoimmunity	N/A		Nil		N/A		Nil			
Pregnancy	N/A		N/A		N/A		N/A		3 in total: 1 elective abortion / 2 stopped trial	

	Avonex®		Betaferon®		Rebif®		Plegridy®		Copaxone®	
DMT/Comparator	Avonex®	Placebo	Betaferon® (8MIUs)	Placebo	Rebif44	Placebo	125µg 2weekly	Placebo	Copaxone 20mg/day	Placebo

Labs	Anaemia		Lymphopaenia		'Asymptomatic ↓in WCC/lymph /neut and ↑LFTs in Rx vs placebo but lessened in Yr 2'		WCC <3		No difference in blood/urine/ECGs btw groups	
	3	1	80	65			7%	1%		
	LFTs/WCC/Plts N		Mild neutropaenia				Lymph <0.8			
			17	4			5%	3%		
			↓Hb / ↓Plts in Rx group but 'clinically insignificant'				No difference ↓Hb / ↓Plts btw groups			
							ALT>5xULN			
							2%	1%		
Neutralising Antibodies (NAbs)	22%	N/A	45%	11%	12.5%	N/A	<1%	<1%	N/A	
					Nab +: no effect on relapse count					
Notes										
	1.5T MRI used Depression scores equal; 1 patient in PBO cohort attempted suicide		Only 5 pts completed 5 yrs of study ARR reduction not significant after 2yrs 1mg=32MIUs MRI 0.15-1.5T: reviewed by 2 radiologists 5yr data: ↓ARR>50% in Nab-ve Rx group for last 2yrs		Dose titration – 20% total dose for 2-4wks, 50% for 2-4wks No difference in depression scores at any time between Rx and PBO		Manufacturer (Biogen®) collected, analysed and contributed to the interpretation of the data 2yr trial		Benefit was greatest if baseline EDSS<2	

	Avonex®		Betaferon®		Rebif®		Plegridy®		Copaxone®	
DMT/Comparator	Avonex®	Placebo	Betaferon® (8MIUs)	Placebo	Rebif44	Placebo	125µg 2weekly	Placebo	Copaxone 20mg/day	Placebo

Table 0-4: Pivotal DMT trials – First-line oral treatments

	Teriflunomide				Dimethyl Fumarate			
DMT/ Comparator	Teriflunomide 14mg	Placebo	Teriflunomide 14mg	Placebo	DMF 240bd	Placebo	DMF 240bd	Placebo
Trial (Year)	TEMSO (2011)[62]		TOWER (2014)[63]		DEFINE (2012)[64]		CONFIRM (2012)[65]	
Trial details								
N (analysed)	358 (333 RRMS)	363 (329 RRMS)	370 (366 RRMS)	388 (366 RRMS)	410	408	359	363
Primary outcome	ARR		ARR		% relapsed by 2yrs		ARR	
Duration	108 wks		108wks (but see Notes)		2yrs		2 yrs (96 weeks)	
Inclusion Criteria	McDonald criteria (2001) EDSS 0-5.5 18-55yrs ≥1 relapse in 1 yr or ≥2relapses in 2yrs RRMS with or without progression		McDonald criteria (2005) EDSS0-5.5 18-55yrs ≥1 relapse in 1 yr OR ≥2 relapses in 2 years – no relapses 30 days before trial		McDonald criteria (2010) EDSS 0-5 18-55yrs ≥1 relapse in 12 months OR ≥1 Gd+ within 6wks of randomisation			
Exclusion Criteria	Relapse within 60 days 'Other systemic diseases' Pregnancy or planning to conceive		Previous cytokine therapy, pregnancy/breast-feeding/planning to conceive Previous IFN or GA within 3months or ever had Natalizumab or other immunosuppressants		'another major disease' LFTs >2x ULN, WCC<3.5, Eosinophils >0.7 Any Rx w monoclonal (except Natalizumab), IFN or GA Rx within 3months		As DEFINE except NO prev Rx with Natalizumab/IVIgs or Plex within 6months	

	Teriflunomide				Dimethyl Fumarate			
DMT/ Comparator	Teriflunomide 14mg	Placebo	Teriflunomide 14mg	Placebo	DMF 240bd	Placebo	DMF 240bd	Placebo
EDSS frequency	After 3 months then 6-monthly thereafter		3-monthly		3-monthly		3-monthly	
MRI frequency	6-monthly		N/A		'MRI cohort'		'MRI cohort'	
					N=184	N=180	N=169	N=167
					6-monthly			
ITT Design	Yes		Yes		?		Yes	
Relapse definition	≥24hrs, stable for 30/7 before and ↑EDSS		≥24hrs AND ↑EDSS by 1 in 2 FSs or 2 in 1 FS (bowel/bladder/cognitive excluded) OR ↑EDSS >0.5		≥24hrs & examination findings ('according to examining neurologist's evaluation')			
Progression definition	↑EDSS ≥1if baseline ≤5.5; ↑EDSS ≥0.5if baseline >5.5 at 3months				↑EDSS≥1 if baseline EDSS ≥1; ↑EDSS ≥1.5 if baseline EDSS 0			
Baseline Characteristics								
Mean age	37.8	38.4	38.2	38.1	38.1	38.5	37.8	36.9
% Female	71	75.8	69	70	72	75	68	69
%White	96.9	98.3	84	82	78	78	85	84
Mean ARR	Relapses in past 1yr (2yrs)		Relapses in past 2yrs		Relapses in past 1yr			
	1.3 (2.2)	1.4 (2.2)	2.1 (1.05)		1.3		1.3	1.4
Mean EDSS	2.67	2.68	2.71	2.69	2.4	2.48	2.6	2.6
Disease duration (Years)	8.7	8.6	8.18	7.64	Time since diagnosis			
					5.6	5.8	4.9	4.8

	Teriflunomide				**Dimethyl Fumarate**			
DMT/ Comparator	Teriflunomide 14mg	Placebo	Teriflunomide 14mg	Placebo	DMF 240bd	Placebo	DMF 240bd	Placebo
MRI Brain T2 data	**Total** lesion vol. (ml)		N/A		Mean No. T2		Mean lesion vol (mm^3)	
	18.08	19.34			47.6	49.2	13,876	14,595
MRI Brain Gd+ data	% with Gd+ lesions		N/A		Mean No. Gd+ lesions		Mean No. Gd+ lesions	
	35.2	38.2			1.2	1.6	2.7	2.7
	Mean No. Gd+ lesions						% with Gd+ lesions	
	1.81	1.66					13	14
Efficacy Outcomes								
Relapses								
ARR (95%CIs)	0.37* (0.31-0.44)	0.54 (0.47-0.62)	0.32* (0.27-0.38)	0.5 (0.43-0.58)	0.19* (0.15-0.23)	0.36 (0.3-0.44)	0.22* (0.18-0.28)	0.4 (0.33-0.49)
Time to 1st relapse	N/A		369 days (282-485)	188 days (142-249)	637	266	25th centile	
							504	210
ARR Reduction	31.5%* (includes non-RRMS patients)		36%* (34.5-37.2)		53%		44%	
% relapse free	56.5 (51-62)	45.6 (50.2 -51)	@ 48 weeks		73*	54	71*	59
			76.3 (71.7-81)	60.6 (55.5-65.6)				

	Teriflunomide				Dimethyl Fumarate			
DMT/ Comparator	Teriflunomide 14mg	Placebo	Teriflunomide 14mg	Placebo	DMF 240bd	Placebo	DMF 240bd	Placebo
Disability								
% progression-free	79.8 (75.3-84.4)	72.7 (67.7-77.7)	@ 48 weeks: 92.2 (89.2-95.1) [n=267]	@ 48 weeks: 85.8 (82.1-89.4) [n=271]	84*	73	87	82
			@ 108 weeks: 84.2 (79.6-88.8) [n=87]	@ 108 weeks: 80.3 (75.9-84.8) [n=83]				
MRI								
Gd+	Mean No. lesions: 0.26*	Mean No. lesions: 1.33	N/A		Mean No. lesions: 0.1* (-0.5 to 0.7)	Mean No. lesions: 1.8 (-2.4 to 6)	Mean No. lesions: 0.5* (-1.2 to 2.2) [N=147]	Mean No. lesions: 2 (-3.6 to 7.6) [N=144]
					% free of Gd+: 93*	% free of Gd+: 62		
T2	%Δ lesion vol.: ↑0.39%* (-6.51 to 7.29)	%Δ lesion vol.: ↑1.67% (-4.8 to 8.14)			No. new/enlarging: 2.6* (2-3.5)	No. new/enlarging: 17 (12.9-22.4)	No. new/enlarging: 5.1* (3.9-6.6) [N=140]	No. new/enlarging: 17.4 (13.3-22.4) [N=139]

	Teriflunomide				Dimethyl Fumarate			
DMT/ Comparator	Teriflunomide 14mg	Placebo	Teriflunomide 14mg	Placebo	DMF 240bd	Placebo	DMF 240bd	Placebo
							%free of new/enlarging	
							27*	12
T1	%Δ lesion vol				N/A		No. new lesions	
	↑0.33%* (-0.68 to 1.34)	↑1.67% (-4.8 to 8.14)					3 (2.3-4) [N=140]	7 (5.3-9.2) [N=139]
Atrophy	BPF Δ from baseline				N/A		N/A	
	-0.003	-0.004						
Safety Outcomes								
% Discontinued Rx	10.9	8.1	16	6	16	13	12	10
					Due to flushing			
					2	<1		
Common AEs with Rx	Diarrhoea (17.9%) Nausea (13.7%) Hair thinning (13.1%) Headache (18.7%) HTN (5%) Skin reactions (11.2%) Pyelonephritis		↑ALT (14%) Hair thinning (13%) Headache (12%) Neutropaenia (9%) Diarrhoea (11%) HTN: 2 serious events in Teri14 group Mean↑BP=2.7/2.2mmHg Peripheral neuropathy (0.8%)		Flushing (38%) Diarrhoea (15%) Abdo pain (10%) Nausea (13%) Lymphopaenia Proteinuria Pruritis ↑AST		Flushing (35%) GI Sx (36%) Proteinuria (8%) [7% placebo group/9% GA group]	

	Teriflunomide				Dimethyl Fumarate			
DMT/ Comparator	Teriflunomide 14mg	Placebo	Teriflunomide 14mg	Placebo	DMF 240bd	Placebo	DMF 240bd	Placebo
% SAE	15.9	12.8	12		18	21	17	22
Malignancy	1 (Cervical)	3	Nil		<1%		Nil in DMF groups [1 placebo, 4 GA]	
Infection	1.6 -2.5% (NS) No serious opportunistic infections		Any infection		Any infection		Any infection	
			44%	51%	64%	65%	56%	50%
			Opportunistic infections (<1%)		Serious infection 2% both groups		No opportunistic infections	
			1 pt intestinal TB	1 pt Hep C/CMV	No opportunistic infections			
Death	No deaths		3 (suicide / septicaemia / RTA)	1	2: RTAs (both in Rx arms)		Nil	
Autoimmunity	Nil		1 pt Teri7mg had ITP – stopped Rx → steroids / Rituximab		Proteinuria		Nil	
					9%	8%		

	Teriflunomide				Dimethyl Fumarate			
DMT/ Comparator	Teriflunomide 14mg	Placebo	Teriflunomide 14mg	Placebo	DMF 240bd	Placebo	DMF 240bd	Placebo
Pregnancy	11: 4 spont abortions (1PBO/3Rx); 6 elective abortions (1 Teri14mg / 5 Teri 7mg) 1 healthy baby @2yrs, after 31 days Teri 14mg in pregnancy		18 14♀: 10 elective abortions; 4 healthy births (1 Teri 14mg) 4♂partners: 1 elective abortion; 3 healthy babies (Teri 7mg)		Nil		Nil	
Labs	Neutropaenia		Neutropaenia		Lymph <0.5		WCC<3	
	<1%	0	3pts (<1%)	0			10%	1%
	↑ALT >1xULN		ALT>1xULN		4%	<1%	Lymph<0.5	
	57.3%	35.9%	55%	39%			5%	<1%
	↑ALT>3xULN		ALT>3xULN		No serious infections in those with lymph<0.5		None stopped Rx due to lymphopaenia	
			8%	6%			ALT>1xULN	
	6.7%	6.7%	Lymph<0.5		ALT>3xULN		47%	41%
			3%	<1%			ALT>3xULN	
			Lymph<0.2		6%	3%	6%	6%
			Nil					
NAbs	N/A				N/A			
Notes								
	TEMSO CIs cross for % relapse free, % with disability progression and MRI outcomes				**DEFINE** SD cross for Gd+ lesions No significant neutropaenia			

	Teriflunomide				Dimethyl Fumarate			
DMT/ Comparator	Teriflunomide 14mg	Placebo	Teriflunomide 14mg	Placebo	DMF 240bd	Placebo	DMF 240bd	Placebo

	Leflunomide used in RA since 1998: 2 PML cases No difference in fatigue questionnaires between groups Bloods every 2wks for 6months then every 6weeks Abdo USS for pancreatitis (associated with Leflunomide) at baseline and every 6/12	LFTs↑ months 1-6 (largest difference with placebo at 4/52) ECGs: No abnormalities detected No renal/liver failure
	TOWER 821 Completed study: placebo=274; Teri7mg=289; Teri14mg=258 Median duration of Rx (weeks): placebo=83; Teri7mg=79; Teri14mg=84 CIs cross for disability No significant difference between placebo and Rx for change in EDSS from baseline to week 48 or changes in physical health summary score, mental health summary score or fatigue score at 48 weeks or at end of trial but *was* significant difference (p=0.02) between mental health summary score and fatigue score (p=0.042) at baseline and end of study between placebo and Teri14mg, but no significant difference at week 48 Most cases of hair thinning resolved on Rx – 2% patients in Teri14mg group stopped Rx due to hair thinning	

	Teriflunomide				Dimethyl Fumarate			
DMT/ Comparator	Teriflunomide 14mg	Placebo	Teriflunomide 14mg	Placebo	DMF 240bd	Placebo	DMF 240bd	Placebo

Table 0-5: Pivotal DMT trials - Second-line treatments

	Fingolimod				**Natalizumab**		**Alemtuzumab**			
DMT/ Comparator	Fingo 0.5	Placebo	Fingo 0.5	Placebo	Nat	Placebo	Alem12	Rebif®	Alem12	Rebif®
Trial (Year)	FREEDOMS I (2010)[66]		FREEDOMS II (2014)[67]		AFFIRM (2006)[68]		CARE-MS I (2012)[69]		CARE-MS II (2012)[70]	
Trial details										
N	425	418	358	355	627	315	376	187	426	202
Primary outcome	ARR		ARR		Relapse rate at 1yr Sustained disability progression at 2yrs		Relapse rate Time to 6-month sustained disability			
Duration	2 yrs		2 yrs		2yrs (120wks)		2 yrs		2 yrs	
Inclusion Criteria	McDonald criteria (2010) EDSS 0-5.5 Age 18-55				McDonald Criteria (2001) EDSS 0-5 ≥1 relapse in past year Age 18-50		Untreated RRMS McDonald criteria (2005) EDSS≤3 Disease duration <5yrs ≥2 relapses in past 2yrs (≥1 relapse in past 1yr) Age 18-50		Relapse despite 1st line Rx McDonald criteria (2005) EDSS≤5.0 Disease duration <10yrs ≥2 relapses in past 2yrs / ≥1 relapse in past 1yr Age 18-55	
Exclusion Criteria	Active infection Recent steroids (30/7) Macular oedema Diabetes				PPMS and SPMS Relapse within 50 days		Progressive MS Previous DMTs Prev monoclonal Abs		As CARE-MS I PLUS Rx with Natalizumab	

	Fingolimod				Natalizumab		Alemtuzumab			
DMT/ Comparator	Fingo 0.5	Placebo	Fingo 0.5	Placebo	Nat	Placebo	Alem12	Rebif®	Alem12	Rebif®

	Immunosuppression ‘clinically significant systemic disease’				Cyclophos or Mitox in previous year IFN, GA, cyclosporine, AZA, MTX, IVIgs in prev 6months (or had IFN or GA for >6/12 at any time)		‘clinically significant autoimmunity other than MS’		within 6/12 [all had prev DMTs¶]	
EDSS frequency	3-monthly									
MRI frequency	6-monthly				12-monthly					
ITT Design?	Yes		Yes		Yes		Yes		Yes	
Relapse definition	≥24hrs, stable for 30/7 and ↑EDSS ≥0.5 (excluding bowel, bladder or cognition)				>24hrs AND signs		>48hrs after 30/7 stability and new signs (↑EDSS)			
Progression definition	Baseline EDSS<5.5: ↑>1 Baseline EDSS≥5.5: ↑>0.5 Confirmed at 3-months outwith relapse				Baseline EDSS≥1: ↑>1 Baseline EDSS 0: ↑>1.5 Confirmed at 3-months outwith relapse		Baseline EDSS≥1: ↑>1 Baseline EDSS 0: ↑>1.5 Confirmed at 6-months			
Baseline Characteristics										
Mean age	36.6	37.2	40.6	40.1	35.6	36.7	33.9	33.2	34.8	35.8
% Female	69.6	71.3	77	81	72	67	65	65	66	65

	Fingolimod				Natalizumab		Alemtuzumab			
DMT/ Comparator	Fingo 0.5	Placebo	Fingo 0.5	Placebo	Nat	Placebo	Alem12	Rebif®	Alem12	Rebif®
%White	N/A		N/A		96	94	94	96	90	93
Mean ARR	No. relapses in past 1yr (2yrs)				Mean relapses in past 1yr					
	1.5 (2.1)	1.4 (2.2)	1.4 (2.2)	1.5 (2.2)	1.53	1.50	1.8	1.8	1.7	1.5
Mean EDSS	2.3	2.5	2.4	2.4	2.3	2.3	2.0	2.0	2.7	2.7
Disease duration	Time since 1st Symptoms				5	6	2.1	2	Time since first clinical event	
	8	8.1	10.4	10.6					4.5	4.7
MRI Brain T2 data	Vol. of lesions (mm^3)				% pts <9 T2 lesions		Vol of lesions (mm^3)			
	6128	6162	5484	5553	5	5	7400	7300	9040	9940
MRI Brain Gd+ data	% with Gd+ lesions				Mean no. Gd+ lesions		Mean no. Gd+ lesions			
	37	38	39	36	2.2	2.0				
			Mean No. Gd+ lesions				2.3	2.1	2.28	2.3
			1.3	1.2						
Efficacy Outcomes										
Relapses										
ARR (95%CIs)	0.18* (0.15-0.22)	0.4 (0.34-0.47)	0.21* (0.17-0.25)	0.4 (0.34-0.48)	ARR @ 1yr		0.18* (0.13-0.23)	0.39 (0.29-0.53)	0.26* (0.21-0.33)	0.52 (0.41-0.66)
					0.26* (0.21-0.32)	0.82 (0.67-0.97)				

	Fingolimod				Natalizumab		Alemtuzumab			
DMT/ Comparator	Fingo 0.5	Placebo	Fingo 0.5	Placebo	Nat	Placebo	Alem12	Rebif®	Alem12	Rebif®
					ARR @ 2yrs					
					0.23*	0.73				
Time to 1st relapse	N/A		Hazard ratio 0.5 (0.4-0.67) delayed time to 1st relapse		N/A		N/A		N/A	
ARR Reduction	55%		48% (47-50)		68%		54% (55-56.6)		50%	
% relapse free	70.4* (66-77.8)	45.6 (40.7-50.6)	71.5* (66.6-76.4)	52.7 (47.2-58.2)	67*	41	77.6*	58.7	65.4*	46.7
Severity	N/A		% Mod/Severe relapses		N/A		N/A		N/A	
			70	77						
Disability										
% progression-free	Confirmed @ 3/12		Confirmed @ 6/12		Hazard ratio 0.46 (0.33-0.64) for risk of progression sustained at 6/12 vs. placebo [54% ↓]		92% [NS]¶	89%	87	80
	82.3 (78.6-86.1)	75.9 (71.7-80.2)	86.2 (77.9-86.4) [NS]	83.2 (77.9-86.4)						
	Confirmed @ 6/12				Cumulative probability of progression				HR 0.58 in favour of Alem: 42% risk reduction*	
	87.5	81			17%*	29%				

	Fingolimod				Natalizumab		Alemtuzumab			
DMT/ Comparator	Fingo 0.5	Placebo	Fingo 0.5	Placebo	Nat	Placebo	Alem12	Rebif®	Alem12	Rebif®
	(84.3-90.7)	(77.1-84.9)								
Change in EDSS	0 (-0.88 to 0.88)	↑0.13 (-0.81 to 1.07)	0.046 [NS]	0.055	N/A		-0.14 [NS]	-0.14	↓0.17* (-0.29 to -0.05	↑0.24 (0.007 to 0.41)
MRI										
Gd+	Mean No. Gd+ lesions (SD)						% with no Gd+ lesions @ 2yrs			
	0.2* (-0.6 to 1)	1.1 (-1.3 to 3.5)	0.4* (-1.44 to 2.24)	1.2 (-1.77 to 4.17)	0.1* (-1.3 to 1.5)	1.2 (-2.7 to 5.1				
	% with no Gd+ @2yrs						93*	81	91*	77
	89.7* [N=369]	65.1 [N=332]	87 [N=269]	65 [N=256]	97*	72				
T2	Mean % Δ Vol.				Mean No. new		Median % Δ Vol.			
	↑10.6* (-92.9-114.1)	↑33.8 (-73.1 to 140.7)	↑13.74*	↑25.06%	1.9*	11	-9.3 (-19.6 to -0.2) [NS]	-6.5 (-20.7 to 2.5)	-1.27 (-12.7 to 7.78) [NS]	-1.73 (-11.1 to 11.39
	Mean no. new or enlarging						% pts with new/enlarging			

	Fingolimod				Natalizumab		Alemtuzumab			
DMT/ Comparator	Fingo 0.5	Placebo	Fingo 0.5	Placebo	Nat	Placebo	Alem12	Rebif®	Alem12	Rebif®
	2.5* (-4.7 to 9.9) [N=370]	9.8 (-3.4 to 23) [N=339]	2.3* (-4.96 to 9.56) [N=264]	8.9 (-4.96 to 22.7) [N=251]			48*	58	46*	68
T1	Mean % Δ Vol.				N/A		N/A		N/A	
	↑8.8 (-67.5 to 85.1) [N=346]	↑50.7 (-337.6 to 439) [N=305]	↑12.6 (-198.47 to 223.67) [N=225] [NS]	↑26.4 (-122.4 to 175.2) [N=209]						
Atrophy	Mean % Δ Brain Vol. (SD)				N/A		Median Δ BPF			
	-0.84* (-2.15 to 0.47) [N=357]	-1.31 (-2.81 to 0.19) [N=331]	-0.86* (-2.08 to 0.36) [N=266]	-1.28 (-0.22 to 2.78) [N=249]			-0.87%* (-1.5 to -0.254)	-1.48% (-2.3 to -0.5)	-0.62%* (-1.3 to 0.006)	-0.81% (-1.5 to 0.2)
Safety Outcomes										
% Discontinued Rx	Due to AEs									
	7.5	7.7	18	10	6	4	1	6	3	7

	Fingolimod				Natalizumab		Alemtuzumab			
DMT/ Comparator	Fingo 0.5	Placebo	Fingo 0.5	Placebo	Nat	Placebo	Alem12	Rebif®	Alem12	Rebif®
	Due to abnormal labs 6.8 / 1.4 Due to lack of efficacy 3.3 / 14.6									
Common AEs with Rx	↑LFTs (15.8%) Influenza infection (12.9%) Back pain (11.8%) Diarrhoea (11.8%) LRTI (9.6%) Herpesvirus infection (8.7%) Dizziness (7.3%) HTN (6.1%) Lymphopaenia (3.5%) Bradycardia (2.1%)		LRTI (11%) Herpesvirus (8%) HTN (9%) ↑ALT (8%) Lymphopaenia (8%) Insomnia (9%) 1st dose 1° AV block (5%) 1st dose 2° AV block (Wenckebach) (4%)		Fatigue (27%) Hypersensitivity reaction (4%)		Infusion reaction (headache, rash, fever) [90%] UTI (17%) Herpesvirus infection (16%) Fatigue (13%) Rash (12%)		Infusion reaction (90%) Fatigue (19%) Chest discomfort (8%) Headache (53%) Dizziness (11%) Rash (44%) Pruritis (15%)	
% SAE	10.1	13.4	15	13	19	24	14	7	13	13
Malignancy	<1% in all groups [BCC 0.9% Fingo0.5 vs 0.2% placebo]		4% (10 pts BCC)	2% (2 pts BCC)	<1% in both groups: 5 in Rx group [3 breast, 1 cervical, 1 melanoma]		2 (1%) thyroid cancers in Alem group (None in Rebif)		2 (BCC, Thyroid)	2 (BCC, AML)

	Fingolimod				Natalizumab		Alemtuzumab			
DMT/ Comparator	Fingo 0.5	Placebo	Fingo 0.5	Placebo	Nat	Placebo	Alem12	Rebif®	Alem12	Rebif®

Infection	69-72% across all groups 1.6-2.6% serious infections		Any infection							
			3%	1%	79% both groups		67%	45%	77%	6%
			VZV infection		Serious infection					
			3%	1%	3.2%	2.6%	2%	1%	4%	1%
Death	0	2	No deaths		2 deaths: both in Rx group [melanoma, alcohol XS]		2 deaths: both in Alem group [RTA, sepsis]		2 deaths in Alem12 group (RTA, aspiration after relapse)	
Autoimmunity	N/A		N/A		N/A		Thyroid		Thyroid	
							18%	6%	16%	5%
							3 pts ITP, 1pt each: haemolytic anaemia, glomerulonephritis, neonatal thyrotoxicosis, pancytopaenia		7 pts ITP, 1 pt glomerulonephritis	
Pregnancy	Nil		Nil		Nil		Nil in Alem group?		Nil	
Labs	Mean 73% ↓Lymph after 1/12 in Rx group		Lymphopaenia (%pts)		↑lymph/mono/eosin in Rx group (not neut)		Lymphopaenia transiently after Alem infusions: B cells recovered within 6 months; T cells after ≈1 year			
	ALT>3xULN		8	0			'Liver toxicity'			
	8.5%	1.7%	↑ALT (%pts)				17%	4%	4%	6%

	Fingolimod				Natalizumab		Alemtuzumab			
DMT/ Comparator	Fingo 0.5	Placebo	Fingo 0.5	Placebo	Nat	Placebo	Alem12	Rebif®	Alem12	Rebif®
			8	2	↑ nucleated red cells in few Rx group					
	ALT>5xULN		Recovery of mean lymphocyte count occurred within 45 days of drug discontinuation		Changes were reversible, without clinical sequelae + returned to baseline approx. 4/12 after last dose					
	1.9%	1%								
NAbs	N/A		N/A		9%	-	29% before 2nd course Alem; 86% 1/12 after 2nd course		29% before 2nd course Alem; 81% 1/12 after 2nd course	
					6% persistent – associated with hypersensitivity and ↓efficacy		Did not influence efficacy or safety			
Notes										
	Bradycardia 8 patients had bradycardia as SAE (7 in Fingolimod groups –		No differences in HRCT between groups; PFTs slightly reduced in Fingolimod groups		MRI 3mm cuts No evidence of 'rebound': 51 Nat patients and 27 Placebo patients		Monthly questionnaires, bloods, urinalysis and microscopy (every 3/12 in Rebif group) TFTs every 3/12		¶ No. previous DMTs tried (mean): Rebif =Alem=1; Drugs used (both groups): Rebif	

	Fingolimod				Natalizumab		Alemtuzumab			
DMT/ Comparator	Fingo 0.5	Placebo	Fingo 0.5	Placebo	Nat	Placebo	Alem12	Rebif®	Alem12	Rebif®
	Fingo0.5=4; Fingo1.25=3) [<1%] 6 were asymptomatic: all continued Fingolimod Total no. bradycardia events: Placebo = 3, Fingo 0.5=9; Fingo1.25=14 6 were symptomatic (dizziness, chest discomfort, palpitations) All resolved within 24hrs; 2 patients required Rx HR decreases started within 2 hrs; nadir after 4-	No difference in LV ejection fraction between groups (echo at baseline, 3 mo, 12mo and 24mo) Bradycardia requiring overnight admission: 2% Fing1.25, none in other groups Most first-dose monitoring events were asymptomatic, did not require Rx and resolved within 24hrs of onset. Max decrease in mean HR in Fingo0.5 group was 8.5bpm Holter ECG findings at 3 months were similar between groups	who stopped Rx returned to baseline disease activity % having relapses after stopping Rx: Nat=29; Placebo=30	Lymphocyte subsets every 3/12 and at 1/12 after Alem infusion Screening for Alem Nabs at baseline, 1,3 and 12 months after each dose IFN Nabs at baseline and 2yrs ¶ authors expected 20% in Rebif group hence perhaps not powered to detect a difference given the low rate in Rebif group %patients MRI and clinically disease free (?at 2yrs): Rebif =27; Alem = 39 Odds ratio = 1.5 (1.2-2.6) [p=0.006]	(≈35%), Avonex (≈25%), Betaferon (≈30%), GA (≈35%), Natalizumab (≈3%), AZA (≈1%) Herpes and fungal infections more common with Alem 1 case VZV requiring hosp admission in Alem12 group (2 in Alem24) Prophylactic Aciclovir reduced herpetic infections (approx. 0.5% with cover,

	Fingolimod				Natalizumab		Alemtuzumab			
DMT/ Comparator	Fingo 0.5	Placebo	Fingo 0.5	Placebo	Nat	Placebo	Alem12	Rebif®	Alem12	Rebif®
	5hrs, with attenuation beginning at 6hrs Max drop in mean resting HR (from baseline) was 8bpm, 5hrs after Fingo0.5 (10bpm 4hrs after Fingo1.25) 1st degree AV block (no. patients): Placebo = 6, Fingo0.5=20 [on day 1 Rx] 2nd degree AV block (Mobitz 1): Placebo=0; Fingo0.5=1, Fing1.25=4 Symptomatic in 1 patient – SOB and palpitations						approx. 2% without) 1 patient pulmonary TB 'No opportunistic infections'			

	Fingolimod				Natalizumab		Alemtuzumab			
DMT/ Comparator	Fingo 0.5	Placebo	Fingo 0.5	Placebo	Nat	Placebo	Alem12	Rebif®	Alem12	Rebif®

	No effects were seen with continued use of Fingolimod **Macular oedema** Macular oedema: 7 Fingo1.25 patients (0 in other groups): 5 occurred within 3 months of Rx; 6 resolved in 1-6months after Rx was discontinued				

Table 0-6: Pivotal DMT trials - Emerging Therapies

	Ocrelizumab				**Cladribine**		**Laquinimod**	
DMT/ Comparator	OCR 600mg 6-monthly	Rebif 44	OCR 600mg 6-monthly	Rebif 44	Cladribine 3.5mg/kg	Placebo	Laquinimod 0.6mg	Placebo
Trial (Year)	2017 (OPERA I)[71]		2017 (OPERA II)[71]		2010 (CLARITY)[72]		2012 (ALLEGRO)[73]	
Trial details								
N (analysed)	410	411	417	418	433	437	550	556
Primary outcome	ARR				ARR		ARR	
Duration	96 weeks				96 weeks		24 months	
Inclusion	Age 18-55 2010 revised McDonald criteria EDSS 0-5.5 at screening 2 clinical relapses in past 2 years OR 1 clinical relapse within 1 year MRI with typical MS changes No neurological worsening for 30 days before screening and baseline visits				McDonald 2001 criteria MRI with typical MS changes ≥1 relapse in past 1 yr EDSS≤5.5 3/12 washout if had DMTs before		Age 18-55 McDonald 2005 criteria EDSS ≤5.5 Disease duration ≥6 months ≥1 relapse in past 1yr or 2 within 2yrs ≥1 Gd+ lesion in previous year	
Exclusion	PPMS Previous B-cell therapy / other immunosuppressants (Alem / Mitox / Teri / Nat >1yr / Cyclophos / MMF / Fingo / DMF) Disease duration > 10 years with EDSS ≤ 2				Failed ≥2 previous DMTs Previous immunosuppressants Abnormal FBC 'Disorder that could compromise immune function' Relapse within 28 days before study		Progressive MS Relapse between screening and baseline Clinically significant or unstable medical or surgical conditions Immunosuppressants in last 6 months Use of IFN/GA in last 2 months Natalizumab ever	
EDSS frequency	3-monthly				3-monthly		3-monthly	

	Ocrelizumab				Cladribine		Laquinimod	
DMT/ Comparator	OCR 600mg 6-monthly	Rebif 44	OCR 600mg 6-monthly	Rebif 44	Cladribine 3.5mg/kg	Placebo	Laquinimod 0.6mg	Placebo
	(not stated but 12wk confirmed disability progression rates reported)							
MRI frequency	6-monthly				6-monthly		12-monthly	
ITT Design	Yes (except NEDA – 'modified ITT')				Yes		Unclear	
Relapse definition	>24hrs and ↑EDSS; stable for 30 days before				>24hrs and ↑EDSS; stable for 30 days before		>48hrs and ↑EDSS; stable for 30 days before	
Progression definition	Increased EDSS Baseline <5.5: ↑ ≥ 1 Baseline >5.5: ↑ ≥ 0.5 Sustained for 3 months				Increased EDSS Baseline 0: ↑ ≥ 1.5 Baseline ≥0.5: ↑ ≥ 1 Sustained for 3 months		Increased EDSS Baseline 0-5: ↑ ≥ 1 Baseline >5.5: ↑ ≥ 0.5 Sustained for 3 months	
Baseline Characteristics								
Mean age	37.1	36.9	37.2	37.4	37.9	38.7	38.9	38.5
% Female	65.9	66.2	65.0	67.0	68.8	65.9	71.1	66.2
%White	N/D		N/D		98.2	98.2	N/D	N/D
Mean ARR	Number of relapses in past 12 months				**N/D**		Relapses in past 12 months	
	1.31	1.33	1.32	1.34			1.2	1.3
Mean EDSS	2.86	2.75	2.78	2.84	2.8	2.9	2.6	2.6
Disease duration	Time since symptoms onset (yrs)				7.9yrs*	8.9yrs	Time since first MS symptom (yrs)	
	6.74	6.25	6.72	6.68			8.7	8.7
T2	Number of T2 lesions (volume, cm^3)				Volume of T2 lesions (mm^3)		Volume of T2 lesions (cm^3)	
	51.04 (10.84)	51.06 (9.74)	49.26 (10.73)	51.01 (10.61)	14828	14287	9.8	9.7

	Ocrelizumab				Cladribine		Laquinimod	
DMT/ Comparator	OCR 600mg 6-monthly	Rebif 44	OCR 600mg 6-monthly	Rebif 44	Cladribine 3.5mg/kg	Placebo	Laquinimod 0.6mg	Placebo
Gd+	% with NO Gd +ve lesions (% with ≥4)				% with lesions		Number of Gd +ve lesions	
	57.5 (14.3)	61.9 (14.0)	61.0 (13.3)	58.6 (14.0)	31.9	29.3		
					Mean number		1.7	2.0
					1.0	0.8		
Efficacy Outcomes								
Relapses								
ARR (95%CIs)	0.16* (0.12-0.2)	0.29 (0.24-0.36)	0.16* (0.12-0.2)	0.29 (0.23-0.36)	0.14* (0.12-0.17)	0.33 (0.29-0.38)	0.3* (±0.02)	0.39 (± 0.03)
Time to 1st relapse	N/D		N/D		13.4* months	4.6 months	N/D	N/D
ARR Reduction	46%		47%		57.6%		N/D	
% relapse free	NEDA (%)		NEDA (%)		79.7*	60.9	62.9*	52.2
	47.9*	29.2	47.5*	25.1				
Severity	N/D		N/D		N/D		N/D	N/D
Disability								
% progression-free	Confirmed at 3 months				Confirmed at 6 months		Confirmed at 3 months	
	92.4*	87.8	90.9*	86.4				
	Confirmed at 6 months				85.7*	79.4	88.9*	84.3
	94.1*	90.5	93.1*	89.5				

	Ocrelizumab				Cladribine		Laquinimod	
DMT/ Comparator	OCR 600mg 6-monthly	Rebif 44	OCR 600mg 6-monthly	Rebif 44	Cladribine 3.5mg/kg	Placebo	Laquinimod 0.6mg	Placebo
Change in EDSS	N/D		N/D		N/D		Mean final EDSS	
							2.68	2.79
MRI								
Gd+	% patients with new lesions				Mean number		Mean number	
	8.3*	30.2	9.8*	36.1			1.33*	2.12
	Mean number of lesions/scan				0.12*	0.91		
	0.02*	0.29	0.02*	0.42				
T2	(New or enlarging) % patients				Mean number		Mean new or enlarged	
	38.3*	61.3	39.1*	62.0	0.38*	1.43	5.03*	7.14
	Mean number of lesions / scan							
	0.32*	1.41	0.33*	1.9				
T1	Mean number of lesions / scan				N/D		N/D	
	0.42*	0.98	0.45*	1.26				
Atrophy	Brain volume change week 24-96 (% change)				N/D		% Change brain vol from baseline	
	-0.57* (-0.66 to -0.49)	-0.74 (-0.83 to -0.65)	-0.64* (-0.73 to -0.54)	-0.75 (-0.85 to -0.65)			-0.87*	-1.3
Patient Reported Outcome / QOL	SF-36 mean score change				N/D		N/D	
	0.04	-0.66	0.33					
Safety Outcomes								
	Due to AEs						Any reason	

	Ocrelizumab				Cladribine		Laquinimod	
DMT/ Comparator	OCR 600mg 6-monthly	Rebif 44	OCR 600mg 6-monthly	Rebif 44	Cladribine 3.5mg/kg	Placebo	Laquinimod 0.6mg	Placebo
% Discontinued Rx	3.2	6.4	3.8	6.0	3.5	2.1	Year 1: 12%	Year 1: 15.2%
							Year 2: 8.5%	Year 2: 7.9%
							Due to AEs: 3%	Due to AEs: 1%
Common AEs with Rx	Infusion reaction (30-37%) Nasopharyngitis (15%) URTI (15%) Headache (11%)				Headache (24%) Lymphopaenia (22%) Nasopharyngitis (14%) URTI (13%) Nausea (10%)		↑ALT (30%) Abdominal pain (6%) Back pain (16%) Cough (8%)	
% SAE	1.2	2.9	1.4	2.9	8.4	6.4	3.9	0.8
Malignancy	3 (0.7%) Ductal breast ca. Renal ca.	1 (0.7%) Lymphoma	1 (0.2%) Melanoma	1 (0.2%) Squamous cell ca	6 (1.4%) 5 benign uterine leiomyoma, 1 cervical ca, 1 melanoma, 1 ovarian ca, 1	0	8 (1.5%) Breast ca (3) Lung ca, BCC, Oesophageal, GBM, Lymphoma (1 each)	6 (1.1%) Ovarian ca (2), Breast ca, BCC, Rectal ca, Prostate ca (1 each)

	Ocrelizumab				Cladribine		Laquinimod	
DMT/ Comparator	OCR 600mg 6-monthly	Rebif 44	OCR 600mg 6-monthly	Rebif 44	Cladribine 3.5mg/kg	Placebo	Laquinimod 0.6mg	Placebo
					pancreatic ca, 1 myelodysplasia (inc 5.25mg/kg group also)			
Infection	Any infection (%)				Serious infection (%)		UTI (7.3%) Sinusitis (5.3%)	UTI (4.5%) Sinusitis (4.5%)
	56.9	54.3	60.2	52.5	2.3	1.6		
	Herpes infections (oral / zoster) (%)				Herpes zoster SAE			
	4.4	3	5.5	3.2	1 patient	0		
Death	0	1 (0.2%) Suicide	1 (0.2%) Suicide	1 (0.2%) Bowel obstruction	2 (0.5%) MI, pancreatic ca	2 (0.5%) Suicide, ICH	0	3 (injury, suicide, pneumonia)
Autoimmunity	Nil		Nil		Nil		Nil	
Pregnancy	Nil documented				1	1	Nil documented	
Labs	Nil documented				Lymphopaenia <0.2 (%)		↑ALT > 3x ULN	
					0.7	0	4.8%	1.6%
					Any thrombocytopaenia (%)		No liver failure / Bilirubinaemia	
					10.9	4.6		
					Neutrophils <1 (%)			

	Ocrelizumab				Cladribine		Laquinimod	
DMT/ Comparator	OCR 600mg 6-monthly	Rebif 44	OCR 600mg 6-monthly	Rebif 44	Cladribine 3.5mg/kg	Placebo	Laquinimod 0.6mg	Placebo
					1.4	1.6		
NAbs	3 patients (0.4%) developed antidrug-binding Abs with Ocrelizumab across both trials 1 developed Nabs 21.3% Rebif patients developed NAbs				Nil documented		Nil documented	
Notes								
	Approx. 25% in each cohort had previous DMTs (IFN>GA>Nat>Fingo>DMF) **As a result of the failure in the statistical hierarchical testing, all the P values for the subsequent secondary efficacy end points, including the change in the SF-36 quality-of-life physical-component summary and the measure of no evidence of disease activity, were considered to be nonconfirmatory: applies to SF-36, NEDA, Brain volume change** Most of the new or newly enlarged lesion activity on T2-weighted MRI in the ocrelizumab groups occurred between baseline and week 24 (Fig. S8 in the Supplementary Appendix). From week 24 to week 48, the number of lesions was 94% lower in the ocrelizumab group than in the interferon				Severe neutropenia (as rated by the investigators) was reported in three patients receiving cladribine (one in the 3.5-mg group and two in the 5.25-mg group), with severe thrombocytopenia and pancytopenia in one of the patients in the latter group, who also had an exacerbation of latent tuberculosis Herpes zoster infections therefore developed in 3.18% of cladribine tablet treated patients		*P values calculated based on adjustments for baseline MRI, disability and/or relapses using regression analyses No significant difference in MSFC scores between LAQ and PBO Safety concerns previously seen with roquinimex, such as serositis, cardiovascular events, and thrombosis, did not emerge as signals in the current study	

	Ocrelizumab				Cladribine		Laquinimod	
DMT/ Comparator	OCR 600mg 6-monthly	Rebif 44	OCR 600mg 6-monthly	Rebif 44	Cladribine 3.5mg/kg	Placebo	Laquinimod 0.6mg	Placebo

	beta1a group in the OPERA I trial and 96% lower in the ocrelizumab group than in the interferon beta1a group in the OPERA II trial. From week 48 to week 96 the number of lesions was 98% lower and 97% lower in the ocrelizumab group than in the interferon beta-1a group in the OPERA I trial and the OPERA II trial, respectively. In the OPERA I trial, a patient treated with ocrelizumab for 1.6 years was hospitalized for a severe genital herpes simplex infection, which resolved with treatment CD19+ cells represent a measure of B-cell counts in anti-CD20–treated patients. The level of CD19+ cells decreased to negligible levels with ocrelizumab treatment by week 2.	experiencing grade 3 or 4 lymphopenia at any time during study compared with development in 1.75% of cladribine tablet treated patients that did not experience grade 3 or 4 lymphopenia during the study The most commonly reported adverse event was lymphocytopenia. There was an inverse correlation between the incidence of infection and a patient's lowest absolute lymphocyte count in the combined cladribine groups. Activation of latent herpes zoster occurred in 20 cladribine-treated patients. One patient who was treated with cladribine had	

	Ocrelizumab				Cladribine		Laquinimod	
DMT/ Comparator	OCR 600mg 6-monthly	Rebif 44	OCR 600mg 6-monthly	Rebif 44	Cladribine 3.5mg/kg	Placebo	Laquinimod 0.6mg	Placebo

		reactivation of latent tuberculosis and died. The use of cladribine may have contributed to this reactivation, and tuberculosis screening measures were subsequently implemented in ongoing clinical trials to rule out latent or active infection before treatment or retreatment.	

	Ocrelizumab				Cladribine		Laquinimod	
DMT/ Comparator	OCR 600mg 6-monthly	Rebif 44	OCR 600mg 6-monthly	Rebif 44	Cladribine 3.5mg/kg	Placebo	Laquinimod 0.6mg	Placebo

Table 0-7: DMT Comparator trials

Trial*	N	Duration	Baseline characteristics				Efficacy Outcomes		
			Age	ARR	EDSS	Disease duration (yrs)	Proportion Relapse-free	EDSS	MRI
Rebif® vs. Avonex® (EVIDENCE)								Proportion 6-month progression	Proportion no new lesions
Rebif®	339	48 weeks	38.3	1	2	4	**62%** (p=0.003)	**12.7%** (HR 0.87, p=0.51)	**63%** (p<0.001)
Avonex®	338		37.4	1	2	4.1	52%	14.5%	43%
Betaferon® vs. Avonex® (INCOMIN)								Proportion 6-month progression	Proportion no new lesions
Betaferon®	96	24 months	38.8	1.52	1.97	5.9	**51%** (p=0.03)	**13%** (p=0.005)	**51%** (p=0.001)
Avonex®	92		34.9	1.38	1.96	6.7	36%	30%	25%
Copaxone® vs Rebif® (REGARD)								Proportion 6-month progression	Proportion no new lesions**
Copaxone®	378	96 weeks	36.8	-	2.33	6.55	62%	8.7% (p=0.117)	31% (p=0.125) [n=230]
Rebif®	386		36.7	-	2.35	5.93	62%	11.7%	38% [n=230]
Copaxone® vs. Betaferon® (BEYOND)								Proportion 3m-month progression	Number of new T2 lesions
Copaxone	448	24 months	35.2	1.6	2.28	5.1	59% (p=0.72)	20% (p=0.68)	**4.6** (p=0.011)
Betaferon® (250µg)	897		35.8	1.6	2.35	5.3	58%	21%	3.3

Trial*	N	Duration	Baseline characteristics				Efficacy Outcomes		
			Age	ARR	EDSS	Disease duration (yrs)	Proportion Relapse-free	EDSS	MRI
Teriflunomide vs Rebif® (TENERE)									
Teriflunomide (14mg)	111	48 weeks	36.8	1.4	2.3	6.6	76.6% (NS)	-	-
Rebif®	104		37.0	1.2	2.0	7.7	84.6%	-	-
DMF vs Copaxone® (CONFIRM)								Proportion 3mo progression	Number of new T2 lesions
DMF (240mg bd)	359	24 months	37.8	1.3	2.6	4.9	71% (p=0.58)	16% (NS)	**5.1** (p=0.007)
Copaxone®	350		36.7	1.4	2.6	4.4	68%	13%	8.0
Fingolimod vs Avonex® (TRANSFORMS)								Proportion 3mo progression	Number of new T2 lesions¶
Fingolimod (0.5mg)	431	12 months	36.7	1.5	2.24	7.5	**82.6** (p<0.001)	6.7% (p=0.25)	**1.5** (p=0.004)
Avonex®	435		36.0	1.5	2.19	7.4	69.3	7.9%	2.6
Natalizumab plus Avonex® vs. Avonex® alone (SENTINEL)								Proportion 3mo progression	Number of new T2 lesions
Natalizumab + Avonex®	589	24 months	38.8	1.44	2.4	7.0	**61%** (p=0.001)	**23%** (p=0.02)	**0.9** (p=0.001)
Avonex®	582		39.1	1.49	2.5	8.0	37%	29%	5.4
Natalizumab plus Copaxone® vs. Copaxone® and PBO (GLANCE)								Median EDSS	Number of new T2 lesions
Natalizumab + Copaxone®	55	24 weeks	40.2	1.4	2.6	8	78% (p=0.658)	2.5	**0.5** (p=0.029)

Trial*	N	Duration	Baseline characteristics				Efficacy Outcomes		
			Age	ARR	EDSS	Disease duration (yrs)	Proportion Relapse-free	EDSS	MRI
PBO + Copaxone®	55		42.5	1.3	2.7	8	73%	2.5	1.3
Alemtuzumab vs Rebif® [Treatment-naïve] (CARE-MS I)								Proportion 6mo progression	Proportion no new T2 lesions
Alemtuzumab	376	24 months	33.9	1.8	2.0	2.1	**77.6%** (p<0.00001)	8% (p=0.22)	**52%** (p=0.04)
Rebif®	187		33.2	1.8	2.0	2.0	58.7%	11%	42%
Alemtuzumab vs Rebif® [Failed 1st DMT] (CARE-MS II)								Proportion 6mo progression	Proportion no new T2 lesions
Alemtuzumab (12mg)	426	24 months	34.8	1.7	2.7	4.5	**65.4%** (p<0.0001)	**13%** (p=0.0084)	**54%** (p<0.0001)
Rebif®	202		35.8	1.5	2.7	4.7	46.7%	20%	32%

*Superior DMT in **bold**

**Less brain volume loss with Copaxone vs. Rebif (-1.24% vs 1.073%, p=0.018)

¶Less brain volume loss with Fingo vs. Avonex (-0.3 vs -0.45, p<0.001)

Table 0-8: Summary of efficacy and safety outcomes from pivotal trials for available DMTs

DMT	ARR Reduction vs PBO	% discontinued Rx
Avonex®	32% [18% (NS) in non-completers]	4%
Betaferon®	33.9%	8.7%
Rebif®	33%	4.9%
Plegridy®	35.5%	4.7%
Copaxone®	29.7%	1.6%
Teriflunomide	31.5-36%	10.9-16%
DMF	44-53%	12-16%
Fingolimod	48-55%	7.5-18%
Natalizumab	68%	6%
Cladribine	57.6%	3.5%
Alemtuzumab	-	1-3%

Emerging therapies and future trial considerations

A number of potential disease-modifying therapies are emerging for both relapsing and progressive forms of MS. The DMTs described so far were developed largely with the aim of immunomodulation or selective immunosuppression, which is reflected in the outcome measures of the pivotal trials – therapies are largely deemed effective due to their impact on relapse rates and MRI inflammatory activity. With time, the focus of therapeutic efficacy has moved back towards preventing established disability, which was in fact the aim of the early IFN trials. The translation into long-term impact from proven (short-term) beneficial effects on measures of inflammatory disease activity remains uncertain, as discussed in our MODERATE study, and so there is a move toward measurable short-term trial outcomes with greater longitudinal relevance.

Essentially, this means better measures of neurodegeneration, the likely mechanism by which long-term disability occurs in MS, though it seems unlikely that this is independent of inflammation and a role for anti-inflammatory treatments will therefore remain. The mechanisms of neurodegeneration in MS and other neurological conditions have not yet been fully elucidated, but loss of neural tissue (as measured radiologically by brain or spinal cord atrophy) likely represents a final common pathway, and so is a potentially useful biomarker to evaluate this. Neural breakdown products such as neurofilaments are another

way in which tissue (axonal) loss can be measured. In terms of clinical outcomes of neurodegeneration, cognitive measures are likely to be important given the multifocal network anatomy involved in higher functions, therefore also providing a final common pathway in grey and white matter loss. Lastly, patient reported outcome measures (PROMs) are now being considered robust enough to base efficacy results upon and of course provide a patient-centred measure of treatment which has long been absent from clinical trials in MS. The move towards this is exemplified by the recent Ocrelizumab studies which have included quality of life measures as a pre-defined outcome for analysis.

With these factors in mind, a role for DMTs targeted at neuroprotection and immunomodulation is suggested but further iterations of established DMTs are likely to be the next available wave of treatments. More selective sphingosine-1-receptor modulators such as Ponesimod [NCT02425644] and Ozanimod [NCT02047734] aim to reduce potential side effects of fingolimod whilst retaining its efficacy. The fully humanised anti-CD20 monoclonal antibody Ofatumumab [NCT02792218] has similar aims as a 'next generation Ocrelizumab'. Novel anti-inflammatory mechanisms have also been explored, such as the anti-CD25 monoclonal antibody Daclizumab which acts through interleukin-2 inhibition, but this has currently been removed from the market because of safety concerns despite having been licensed initially[74,75]. A number of neuroprotective agents have been studied, with varied results, but none have yet reached licensed therapy status[76]. A specific approach is promotion of remyelination, Opicinumab being an example which is currently being tested in a Phase II/III study [NCT03222973]. Opicinumab actively binds Nogo receptor-interacting protein-1 (LINGO-1) which otherwise negatively regulates myelination and inhibits neurite outgrowth. A Phase II study did not meet its primary endpoint but this may also reflect the challenge of demonstrating effects on neurodegeneration given the likely longer trial durations required to do so[77]. Additionally, phenytoin, through its sodium-channel stabilisation effects, has been shown to reduce retinal neural loss when used early in optic neuritis[78] but further study is needed to confirm the utility of this in larger cohorts.

Indeed, the repurposing of drugs licensed for other indications has become a strategy for drug discovery in MS, obviating the need for time-consuming and expensive development programmes essentially available only to large

multinational pharmaceutical companies. Minocycline[79] and clemastine[80] are examples of this for RRMS but this approach has been used for progressive MS given the accessibility of these to smaller institutions where the financial burden of drug discovery would otherwise preclude this.

The focus of this thesis is on RRMS and a detailed discussion on treatment of progressive MS is beyond its scope. That said, the development of DMTs for progressive forms of MS, both SPMS and PPMS, bears mentioning given the therapeutic nihilism that has existed until now. The FDA has approved Ocrelizumab for use in PPMS following a successful demonstration of benefit in a placebo-controlled Phase III study[81]. This is the first licensed therapy for PPMS. Similarly, Siponimod (another S1P-modulator) has been shown efficacious in SPMS[82] and high dose Biotin, an endogenous co-factor ('Vitamin H'), also appears promising and remains under evaluation as a neuroprotective agent thought to act through promotion of remyelination, amongst other mechanisms[83]. The effectiveness of immunomodulatory therapy serves as evidence for the hypothesis that the progressive forms of MS are simply a different stage of the disease rather than a separate entity to RRMS, and that DMTs are likely to be effective for both if outcome measures are appropriately applied over the correct duration of time[84].

Lastly, the focus on MS in isolation has recently been questioned by evidence supporting the role of co-morbidities in detrimental outcomes for people with MS[85]. Clearly this is particularly relevant for progressive phenotypes due to advanced age but is also relevant in RRMS where present. The trials presented here generally exclude significant co-morbidities in order to demonstrate efficacy in the absence of significant confounders but in real-world practice this does not apply. This, firstly, calls into question the applicability of the stated results to populations who were not included in the trials and, secondly, the unmeasured impact of non-MS diagnoses on their outcomes. Whilst younger patients with RRMS may have few, if any, co-morbidities at diagnosis these will inevitably accumulate as per the usual epidemiology of their location, genetic background and lifestyle factors common to all in society. Remaining cognisant of other co-morbidities is therefore vital is considering the role of MS as perhaps only one facet in a holistic care model.

In conclusion, the natural history of RRMS is well established from work in the pre-treatment era and there are now numerous treatments with demonstrable disease-modifying capabilities which have been studied in large randomised controlled trials. There remain issues in extrapolating natural history data to the modern era, not least because of changing definitions of MS over time, and one must be aware of the limitations of trial data. However, the growing number of treatment options for patients with MS, including progressive forms, is welcomed and provides choice to clinicians and patients which was not available in the past. The following chapter reviews the literature on the use of DMTs in practice and summarises some of our own experience with recently-introduced oral agents.

Chapter 1: DMT initiation and escalation in RRMS

Introduction

Literature Review

At present, there are over ten licensed disease-modifying therapies (DMTs) for RRMS, with more expected in the near future, as described earlier. This chapter evaluates the literature regarding DMT initiation and escalation in general and the use of Dimethyl Fumarate (DMF) and Fingolimod in our centre in particular as examples of the role of real-world evidence supplementing randomised trials.

DMT initiation

Evidence of asymptomatic radiological and pathological CNS changes suggest that by the time first symptoms occur in MS the disease process is often already well established. There is a spectrum of disease from asymptomatic microscopic pathological abnormalities to macroscopic radiological changes (RIS) and symptomatic clinical events which can be isolated (CIS), recurrent (RRMS) or progressive (PPMS). The best point at which to initiate DMTs along this timeline must therefore be considered, in the knowledge that not all those with asymptomatic lesions will be affected and that the prognosis is highly variable even in those who develop symptoms. There remains significant variation in the practice of clinicians and the guidance of national advisory bodies worldwide as to when DMTs should be instituted, or offered at least, and the evidence often appears contradictory. Similar to the apparent dissociation between inflammatory and degenerative processes from an immunopathological view in MS, there appears to be disparity between short-term beneficial effects of DMT initiation early in the disease course and long-term outcomes. This evidence largely comes from trials of IFNs however, and so may have limited applicability to newer DMTs that have demonstrably more powerful anti-inflammatory effects. Additionally, the role of induction therapies at an early stage shows promise in modulating both inflammatory and degenerative activity but their relatively recent development limits conclusions about long-term outcomes.

Notably, all of the studies cited to demonstrate benefits of early treatment, or not, suffer from methodological limitations, not least the use of the EDSS as an outcome measure. This largely relies on motor disability to demonstrate

symptomatic worsening, when non-motor symptoms can often be just as disabling.

Timing of treatment in Multiple Sclerosis

By definition, patients must have had clinical symptoms to meet diagnostic criteria for MS[86]. Further, these must be 'typical of an acute inflammatory demyelinating event in the CNS' with non-focal, but arguably CNS-related, symptoms such as fatigue or depression not included. For a CIS, a single clinical demyelinating event has occurred; this can also be the case for RRMS if dissemination in time and space can be otherwise demonstrated. With the advent of MRI for investigation of a wide range of neurological presentations, radiological changes typical of MS, including dissemination in time and space, have been seen in patients with no clinical history of typical demyelinating events. This has been termed the Radiologically Isolated Syndrome (RIS)[87]. This is currently the earliest in vivo manifestation of subclinical demyelination available to clinicians and researchers. Importantly, this does appear to be a prelude to MS development as the majority develop further radiological disease activity within 2-5 years and around a third have a clinical demyelinating event during this time[88]. There have been no attempts at systematic treatment trials in RIS, but rather single-figure cases of treated patients included in MS study cohorts with off-label DMT use. There is therefore no robust data on treatment outcomes in RIS and the focus remains on delineating its natural history.

Given the relatively recent discovery of RIS and uncertainty as to its natural history, CIS is currently the earliest stage of CNS demyelination where the utility of DMTs has been evaluated in a systematic way. There have been a number of clinical trials of DMTs in CIS, as described below, but there are other data which support the consideration of treating demyelination at the earliest opportunity[89] and these are considered first.

'Epitope spreading' describes the propagation of immunological self-recognition with recurrent exposure of a single antigenic stimulus. This was described in the animal model of MS, Extrinsic allergic encephalomyelitis (EAE), and purports to explain, in part at least, the increasing complexity of the immune response with time. Using T-cells activated against a single epitope in proteolipid protein

(PLP), an abundant myelin component, the EAE mouse model was created. Yet, after recurrent relapses in the animal, T-cells reactive against another epitope were demonstrated[90]. This appears to be a fundamental immunological phenomenon and suggests that autoimmunity begets autoimmunity such that the host response multiplies its defence against a single antigen to include other (potential) antigens, increasing the likelihood of erroneous self-attack. It follows, therefore, that if the immune response can be halted before the propagation of self-recognising epitope spreading, perhaps this will limit the damage caused by any multiplied response. The counter to this is the known limitations of animal models in general, and EAE in particular, in replicating the human situation.

Early treatment can also be considered in order to prevent the axonal loss thought to underlie persistent disability in MS. Since the first pathological diagnosis of multiple sclerosis, axonal damage has been known to occur[38] and this has been confirmed in modern studies, where axonal transection was seen in lesions from patients with disease duration from two weeks to 27 years[91]. Of course, the availability of biopsy tissue marks these cases out as unusual, but it is generally accepted that this is a feature of all MS lesions. In early disease, compensatory factors may prevent irreversible axonal loss, particularly via the integrity of the oligodendrocyte-axonal unit. With recurrent injury however, compensatory mechanisms such as remyelination may fail. This, along with the apparent reduction in axonal loss over the course of the disease[92], suggests that early treatment before irreversible axonal loss occurs and during the period of maximal loss may be the best approach to reduce long-term disability.

Imaging studies provide in vivo evidence of the impact of MS even at the earliest stages and, hence, potential treatment targets. MRI correlates of volume have been used to estimate atrophy within CNS structures over time and are thought to reflect the sum total of both neurons and axons - lower volumes result from the loss of both. Whilst volume loss is not necessarily apparent in all patients with CIS, it occurs at a higher rate in those who go on to develop RRMS, as measured by ventricular enlargement[93] and corpus callosum atrophy[94]. Similarly, comparing patients with MS to age-matched healthy controls, there are significant reductions of brain and spinal cord volume measured on MRI[95]. Additionally, those with SPMS (and longer disease duration) have significantly

lower brain and spinal cord volumes than those with RRMS, suggesting the atrophy continues to progress over the disease course. Further, it has been established that MRI measures of atrophy, particularly spinal cord grey matter atrophy[96], correlate with measures of disability[97,95] and that the treatment effect of DMTs on disability independently correlates with their effects on MRI brain atrophy measures[98]. These data suggest that CNS atrophy occurs at the earliest stage of MS and could be used as a treatment target biomarker to assess the efficacy of DMTs in relation to disability progression. It should be remembered, however, that the 'disability' with which atrophy was correlated by Sormani et al.[98] was the short-term disability progression at 3- or 6-months from large clinical trials which has since been demonstrated as frequently reversible[35] and perhaps relapse-related, rather than indicative of relentless disability worsening.

The concept of a 'window of opportunity' for treatment, that there is a timepoint after which treatment will be less or ineffective in MS, derives from the use of Alemtuzumab in patients with RRMS and SPMS. The first use of Alemtuzumab in MS included 58 patients with relapsing-remitting and secondary progressive disease. This was an open-label study conducted between 1991 and 2002 with a mean follow-up of 29 months[99]. In patients with established secondary progressive disease (n=36), a significantly reduced relapse rate [0.7/yr to 0.001/yr ($p<0.001$)] and absence of new lesions on MRI was accompanied by progressively worsening disability of the group over the study period, as measured by the Expanded Disability Status Scale (EDSS). When used in patients with RRMS (n=22), in whom licensed treatment had failed or who had a high early relapse rate suggesting poor prognosis, the annualised relapse rate (ARR) improved by 94% in comparison to the year before treatment, and, unlike the SPMS cohort, EDSS improved - by a mean of 1.2 points 2 years after a five-day course of 20mg Alemtuzumab/day intravenously. Following this, a 'window of opportunity' to slow disease progression was suggested in order to explain the difference in the outcomes between patients with early versus established disease, despite similar reductions in inflammatory clinicoradiological activity. The presence of secondary axonal damage was the postulated mechanism for the lack of effectiveness in patients with progressive disease. This hypothesis, with

rational clinical and pathological evidence to support it, provides further weight to the argument that early treatment is likely more beneficial.

What is not clear, however, is when this window of opportunity 'closes'. Those with established SPMS and resultant disability did not appear to benefit, but at what point does SPMS become 'established' and 'untreatable'? There are no accepted clinical criteria and there is no diagnostic test that demarcates when a patient enters the secondary progressive phase. A standardised definition has been recently proposed[100]. A pragmatic consensus opinion generally defines SPMS as the time-point when disability worsens in the absence of (explanatory) relapses and usually the disability progression must be confirmed after 1 year. Both the relapsing-remitting and progressive phases can be further categorised as 'active' or not and 'progressive' or not, depending on the occurrence of subsequent clinico-radiological disease activity or worsening disability on an annual basis[101]. Providing more clarity on the definitions of the dynamic disease course of MS may assist in more robust 'cut-offs' for the therapeutic window and ensure those most likely to benefit receive appropriate treatment at the right time. Despite these clinical findings, however, patients with progressive MS continue to have evidence of CNS inflammatory activity[102], suggesting there remains an inflammatory target in the progressive phase which may be amenable to treatment. The uncertainties underlying this may, however, contribute to therapeutic inertia in MS whereby potentially disease-modifying therapies are not instituted or considered despite evidence of disease worsening.

Trials of DMTs in early MS and predictors of treatment response

Trials of DMTs at the earliest clinical stage of MS, Clinically Isolated Syndromes (CIS), suggest there is a short-term benefit but this does not appear sustained in long-term extension studies. There have been 4 large trials of DMTs in CIS[103–106]. These have all shown significant reductions in conversion to MS from CIS in the treated cohorts in comparison to placebo. Between 43 and 50% of the untreated CIS cohort converted to MS in comparison to between 25 and 35% of the treated patients over 2-3 years.

Kappos et al. describe benefits of early treatment in terms of disability levels, relapse rates and MRI burden of disease after 8 years of follow-up comparing

those commenced on IFN (Rebif®) immediately or delayed (after 2 years) as part of a RCT, but also outline the issues inherent in long-term follow-up studies to which theirs and others are subject[107]:

1) Patient ascertainment – those remaining under long-term follow-up will tend to be tolerating the treatment and perceiving it as beneficial more so than those who are lost to follow-up as they choose to stop treatment, pursue alternatives or simply cannot participate in the study requirements due to accumulated disability i.e. these studies will favour treatment-responders
2) Lack of long-term parallel placebo control group
3) Changing (improved) MRI technologies which may affect data
4) Retrospective assessment of relapses and adverse events may result in recall bias
5) Difficulties confirming EDSS progression
6) Unblinded assessment of patients no longer on study treatment
7) Treatment interruptions and conversion to non-study medications

In terms of disability outcomes, they found that 23.9% (32/134) of those initially randomised to Rebif® 44mcg TIW reached EDSS 4.0 by 8 years in comparison to 27.6% (37/134) in the delayed treatment group. They could not make comparisons of SPMS development as this was not collected in the original study.

Goodin et al[108] found that long-term physical and cognitive outcomes were not significantly affected by whether patients were randomised to Betaferon® or not within the first 2 years. Assessing patients 16 years after their involvement in a pivotal IFN trial, the authors sought to identify which factors in that early course of their disease related to long-term outcomes. Reflecting that this study is similarly subject to the methodological flaws outlined above, only 260 (69.8%) of those involved in the original trial were included in this analysis. Those who participated in the long-term follow-up were largely comparable to those who did not however, other than the proportion that received active treatment (IFNß-1b 250µg) during the 2-year trial was higher in those included in the long-term follow-up in comparison to those who were not (37% vs 25%, p=0.0178). Using univariate regression, a number of variables were assessed for their correlation with physical (EDSS or SPMS conversion) and cognitive (summary of

PASAT, SDMT, CVLT-II, COWAT and D-KEFS tests) outcomes. Physical outcomes were statistically significantly correlated with baseline EDSS, relapses during the trial and EDSS change over the course of the trial. However, the correlations were fairly weak (R^2 0.22 for baseline EDSS and 0.12 for on-trial ARR). Similarly, baseline EDSS, baseline MRI T2 lesion volume, third ventricular width (baseline and change during trial) and premorbid IQ had weak (but statistically significant) correlation with cognitive outcomes at 16 years. Notably, being on treatment during the 2-year trial period did not correlate significantly with either physical or cognitive outcomes at 16 years. Similarly, the on-trial change in T2 MRI brain lesion volume, a common short-term outcome measure, did not correlate with either physical or cognitive outcomes. The authors conclude that long-term outcome in MS may be largely determined early in the disease course and that baseline characteristics, before treatment, appear more important than the treatment thereafter. Again, this may suggest that treating those with the most severe disease earlier is optimal and that patients presenting with more established disability will benefit less from treatment in the long-term.

Interestingly, the same group evaluated this cohort at 21 years with regards to overall survival[109]. In this study, they were able to determine the relevant data for 366 (98.4%) of the original trial cohort (121 original PBO arm; 245 original IFN arm) and found that all-cause mortality was significantly lower in those included in the original treatment arm (mean 21.1 years previously) in comparison to those randomised to placebo. This was in spite of the fact that all were then offered open-label treatment and did not differ systematically (although the treatments used after the 2-year trial are not presented). A total of 81 (22.1%) deaths occurred over the follow-up period, 37 who had originally been in the placebo arm (30.6%) and 44 who had been in the treatment arm (18%). Those in the high-dose treatment arm had a reduced hazard of death during follow-up in comparison to the placebo group (HR=0.532, 95%CIs 0.314-0.902, p=0.0173) hence there was almost a 50% relative reduction in mortality in the treatment group. The groups were randomised, hence, whilst co-morbidities and other confounders for all-cause mortality are not presented, it would be assumed they were balanced on such measures. Additionally, of those whose cause of death is reported, the majority were considered MS-related, raising the possibility that early treatment may have prevented these. This finding has not been evaluated

in a prospective, systematic way, however, therefore it is difficult to form firm conclusions regarding the impact of early treatment on longer-term mortality from this analysis alone, but a recent study of the Danish MS register also found evidence for a shortened time to death for patients starting DMTs 2 years after their first symptom in comparison to those starting sooner, after over 10 years of follow-up[110].

In an attempt to overcome some of the systematic methodological flaws common to long-term follow-up studies, Trojano et al. used statistical matching in order to compare contemporary treated and untreated cohorts followed for a median of 5.7 years and found significant benefits in the treated group[111]. Patients with RRMS (Poser and McDonald criteria) were recruited from 2 Italian centres and treatment with IFN assigned at the discretion of the treating physician and patient as part of usual clinical practice. Those treated with IFN (n=1103) were then compared with those not treated (n=401). Obviously, the cohorts differed due to indication bias, amongst other variables, hence were not comparable. The cohorts were therefore matched statistically by including a propensity score within Cox proportional hazards regression models. Whilst 23% of the untreated cohort did not receive treatment because their disease was considered too mild (no relapses in past 2 years and EDSS<3), a number of other factors were relevant to this group: refused treatment (19%), planning pregnancy (15%), discontinuation of DMTs within 3-6 weeks due to adverse events and significant concomitant disease preventing the use of IFN (23%). Time to SPMS and disability milestones EDSS 4 and 6 were the endpoints compared in the matched samples. The cohort size was reduced in the matching process (as not all participants had a comparable match), such that n=1328 for the SPMS endpoint, 1246 for EDSS 4 and 1378 for EDSS 6. Using these matched cohorts, there was a significant reduction in the risk of developing SPMS [HR 0.38 (95%CIs 0.24-0.58) p<0.0001] in the IFN-treated group in comparison to the untreated group. For EDSS 4 and 6, there was also a lower risk in reaching these endpoints in the treated versus untreated patients [HR 0.7 (0.53-0.94), p=0.0174 and HR 0.6 (0.38-0.95), p=0.0304 respectively]. The percentage of patients that reached SPMS after 7 years of follow-up was 20.2% for the untreated group versus 8% for IFN-treated patients. Similarly, in terms of proportions, 28% of untreated patients had reached EDSS 4 by 7 years and 12.4% EDSS 6; for IFN-treated

patients, 20.5% and 7.7% reached these endpoints respectively. Sensitivity analyses were undertaken which suggested only an unmeasured confounder between the groups with a hazard ratio of ≥2 would be required to lose the observed significant effect of IFN on the measured outcomes. The authors conclude that this study suggests that IFN-β delays the inevitable and irreversible clinical worsening seen in pre-treatment era natural history cohorts, which the untreated group was comparable with in terms of outcomes. They concede the usual biases associated with observational studies but contend that the statistical matching process at least accounts for overt confounders and the sensitivity analyses, whilst suggesting an unknown confounder could invalidate the results, do not imply a confounder necessarily exists. The study was heavily criticised as it was thought to specifically suffer from immortal time bias (described earlier), and when it was reanalysed taking this into account the positive results did not persist.

These data provide some support for the early use of DMTs in RRMS, but once the decision to embark upon DMTs has been made, a number of questions naturally follow. Firstly, with numerous licensed therapies, which one to choose can be somewhat daunting for both patients and clinicians given the differing properties of each in terms of efficacy, safety, administration and monitoring. Cost is certainly a consideration for patients in fee-paying healthcare services and obviously impacts upon approvals in the UK, but is not considered in detail here. The wealth of data from large, high-quality, randomised trials is certainly empowering, but can also be potentially overwhelming in treatment decisions for patients and non-specialists. The lack of head-to-head trials for most of the available DMTs results in an inability to derive fully evidence-based conclusions on relative safety and efficacy outcomes between many of the options.

Certain DMTs have pronounced immunomodulatory effects such that their use can be considered to have long-term effects even after treatment is discontinued. The majority of licensed therapies, however, are seen as maintenance therapies that lose their efficacy when stopped, albeit some appear more effective in reducing disease activity than others. These two classes of DMT belie the treatment paradigms of induction and escalation in MS respectively. The decision to use a high-efficacy broad immunomodulator at the outset is tempered by the higher risks associated with such treatment and the

difficulty in accurately prognosticating long-term outcomes in MS at an early stage and, hence, who is likely to benefit most. Equally, high-efficacy maintenance therapies have significant associated risks and the escalation approach demands failure of initial treatment(s), exposing the patient to a sequence of immunomodulating therapies, the long-term safety of which is not established. Further DMT use may be necessary in patients despite an induction approach and of course the same concerns apply but it appears the frequency of this is significantly lower.

Escalation and Switching

It is clear that MS disease activity continues in a significant proportion of patients after DMTs are started and deciding if and when to change treatment remains a common clinical problem. This is largely informed by perceived benefits of reducing overt clinical or radiological disease activity in the short-term although the level of disease activity which is reasonable to tolerate (if any), what therapy should be offered as an alternative, and the impact on long-term outcomes remains controversial.

DMT therapeutic paradigms

Induction

The concepts of induction and escalation in MS therapeutics result from the growing number of DMT options over recent years, with differing mechanisms of action and efficacy. Thus, when IFN and GA were the only licensed options, there were obvious limits in the therapeutic strategy. Freedman defined induction as moving either transiently or completely to a 2nd-line drug before ever attempting first-line therapy[112], whilst Rieckmann described it in terms of the goal of 'resetting' the immune system to avoid epitope spreading and prevent early structural damage[113]. The latter mechanistic description seems most appropriate as 2nd-line therapies are often designated so due to safety, and possibly economic, issues rather than their mechanism of action: hence, Fingolimod and Natalizumab are considered 2nd-line therapies but not with the aim of inducing a permanent alteration of the immune system. Additionally, the understood mechanisms of action of these agents and the apparent reversibility of their pharmacodynamic effects on discontinuation suggest there is no

fundamental change to the immune system with their use. In contrast, Mitoxantrone impacts upon DNA and hence protein synthesis upstream of circulating immune cells, suggesting a basic alteration in immune (and other) cell-lines.

Alemtuzumab is perhaps the clearest example of an induction agent given its demonstrable CD52-mediated depletion of lymphocytes and resultant subsequent switch from a pro-inflammatory Th1 lymphocytic phenotype to anti-inflammatory Th2 after treatment[114]. This measurable re-population of lymphocytes over years after a short course of Alemtuzumab treatment implies a fundamental change in the immune repertoire, but whether this is maintained thereafter is not known. So-called induction agents can be used first-line if induction is the goal and this would appear to be the logical time for greatest benefit, but the inherent adverse safety profile associated with broad immunosuppression limits their use in this way. For this reason, induction is generally reserved for those with active, aggressive disease where the balance of risks can be considered to favour treatment risks over disease risks. The difficulty remains, however, on an individual patient basis, in predicting whether their early active disease will undoubtedly result in long-term disability. Equally, reserving induction therapies for those with highly active disease implies reassurance for those with less disease activity who may actually have comparatively worse long-term outcomes, perceived or actual, depending on their lifestyle, the future characteristics of their disease or their goals.

There is precedent in the induction approach in other conditions. Use of Cyclophosphamide followed by maintenance steroids is the standard approach in ANCA-associated systemic vasculitis[115] and bone marrow transplant in conjunction with cytotoxic chemotherapy has resulted in drastic disease-free survival benefits in haematological malignancies which were previously fatal[116]. Of course, MS is not a life-threatening illness hence the aggressive approach in other conditions is not comparable, given the treatment risks, but the model is applicable.

The focus on MS as a T-cell dependent autoimmune condition has been challenged in the last decade, not least by the efficacy of Alemtuzumab which depletes B- and T-lymphocytes. This recognition has led to investigation of a

number of B-cell depleting agents in recent years, with positive results in both relapsing and progressive MS phenotypes[71,117-119]. Rituximab and its humanised analogues Ocrelizumab and Ofatumumab are anti-CD20 monoclonal antibodies which target circulating B-cells but, importantly, do not affect long-term memory and early progenitor cells and so have more predictable and less widespread deleterious effects: this seems borne out by favourable safety profiles from trial experience thus far. It is proposed that the benefits of B-cell depletion include T-cell effects given the loss of antigen presentation and secondary effects, rather than purely through B-cell loss, retaining the role of T-cells in the immunopathogenesis of MS. Whilst there is lymphocyte depletion however, typical of an induction agent, this does not appear sustained in the same way as Alemtuzumab's effects, so long-term dosing is advocated for these agents. This is perhaps a benefit, given the ability to discontinue treatment and, presumably, reverse or reduce deleterious effects, but requires long-term financial expenditure. In this sense, it is proposed that B-cell depleting agents are used as a maintenance therapy despite their induction qualities.

The common theme to these agents is their ability to deplete immune cells, resulting in necessary re-population, rather than simply reducing cell trafficking or modulating their function. This non-targeted approach clearly has undesirable wider effects and has limited the use of the induction approach from a practical point of view as much as any theoretical debate over MS treatment paradigms. Alemtuzumab and Cladribine are licensed induction agents in the UK presently but only Alemtuzumab can be used as a first-line therapy although this is not the case globally. The expected introduction of DMTs with induction properties yet more favourable safety profiles is likely to result in more widespread use of this approach.

Escalation

The current standard treatment paradigm for RRMS remains escalation. In this approach, the safest treatment is used first and switching to higher risk DMTs is only considered if this fails. The higher risk DMTs generally equate to higher efficacy but the lack of head-to-head randomised trials prevents this conclusion being definitive. From the available evidence where comparator trials have been undertaken, however, some conclusions can be drawn as outlined earlier. Alemtuzumab was more effective than Rebif®, whether patients were

treatment-naïve[69] or had failed previous treatments[120]. The addition of Natalizumab to Avonex®[121] or Copaxone®[122] results in better efficacy outcomes than these agents as monotherapy, although benefits were minimal in the latter. Fingolimod had better efficacy outcomes in direct comparison with Avonex®[123] and there was a trend, but non-significant, to Dimethyl Fumarate comparing favourably with Copaxone®[65]. There was no demonstrable difference between Teriflunomide and Rebif® in terms of efficacy[124] but Rebif®[47] and Betaseron®[125] were superior to Avonex®. Copaxone® was equivalent to both Rebif®[126] and Betaseron®[127].

Blinding is problematic in these trials, given the varied modes and frequency of administration. Indeed, the issue with a number of these studies is the lack of double-blinding and often short follow-up periods, again, making definitive conclusions difficult. That said, the escalation paradigm presupposes superiority of some DMTs over others in this way. Given the lack of randomised and double-blinded trials with adequate follow-up, post-hoc re-analyses of the data have been undertaken, for example, as network meta-analyses[52,53]. Of course, these are based on the flawed data described and placebo-controlled trials, meaning that conclusions are not equivalent to a prospective, multi-arm randomised, blinded trial. However, a 3-tier efficacy ladder is generally accepted, with the monoclonal antibodies Natalizumab and Alemtuzumab likely most efficacious and the injectable therapies (IFN/GA – 'BRACE' – an acronym of their trade-names used above) least effective. Fingolimod is considered more effective than Dimethyl Fumarate and Teriflunomide, which are considered either equivalent or Dimethyl Fumarate possibly superior. However, Teriflunomide is the only oral DMT associated with reduced risk of sustained progression of disability (confirmed at 3 months) in both its pivotal placebo-controlled trials[62,63]. This applies to the hazard ratio of sustained disability progression [0.7 (95%CIs 0.51-0.97, p=0.03)] and time to sustained progression [0.68 (0.47-1.0, p=0.0442)] but the proportion of patients free from sustained progression, as reported in the other oral DMT trials, was not significantly different to placebo, as the 95% CIs cross.

From the head-to-head studies cited, this 'order' of DMT efficacy seems credible, with all agents comparing favourably to Avonex® where tested; Copaxone®, Rebif®, Betaseron® and Teriflunomide appearing equivalent to each

other; Dimethyl Fumarate trending towards superiority against Copaxone® and Alemtuzumab clearly superior to Rebif® (in an unblinded trial). The magnitude of Fingolimod's efficacy over Avonex® compares favourably with Avonex's comparisons to other injectable treatments, again, suggesting a degree of superiority above these. Natalizumab is the slight outlier, given its lack of direct comparison to the other DMTs or trial evidence of use in patients failing previous treatment. Again, though, the magnitude of its efficacy over placebo belies its position as a higher-efficacy agent in comparison to placebo-controlled trials of the other DMTs. Indeed, this order is borne out in the network meta-analyses cited above and the real-world observational studies discussed below.

Of course, a major flaw in all of these apparent comparisons is the outcome measure used to compare them. The above descriptions are largely based on relapse rates, which have previously been discussed as of potentially limited value as an outcome measure, particularly if long-term disability prevention is the goal. Short-term (3- or 6-month) disability worsening and MRI lesions and/or atrophy are alternatives but also have flaws in their predictive value of long-term disability. The 'efficacy ladder' is therefore based on short-term outcomes and weighted towards relapse rates as this has been the primary outcome for the majority of DMT trials. Quality of Life and other PROMs are not assessed in a way which permits comparison of DMTs based on these important outcomes either, for example. In TENERE, however, scores on the Treatment Satisfaction Questionnaire for Medication (TSQM) were a secondary endpoint, with significantly higher scores for Teriflunomide over Rebif® (68.82 vs 60.98, p=0.02) despite no demonstrable efficacy difference[124]. The role of PROMs as an outcome measure is likely to increase over time, particularly if efficacy and safety are comparable between agents: this would require a paradigm-shift in MS trials, however, and necessitate much improved understanding of the interpretation and limitations of such measures, including their validity and precision, within the global MS community. Certainly, robust (ideally short-term) outcome measures which predict long-term disability accurately remain elusive but are necessary for future DMT comparator studies to allow accurate ordering of their efficacy. Serum, CSF and MRI biomarkers are postulated as potential outcome measures to address this, but are not yet validated in this way.

Safety considerations are of course paramount in the minds of clinicians and patients when comparing DMT options. The majority of adverse events are usually recognised during Phase 2 and 3 studies, permitting their anticipation (and therefore mitigation) once drugs are in clinical use. However, the development of PML in 2 patients involved in the SENTINEL study (combining Natalizumab and Avonex®), after the study in one case, raised major concerns about the risks of DMTs which may not be apparent from a 1-2 year trial. These events led to the withdrawal of Natalizumab from the market and its re-introduction only after a strict risk mitigation scheme was developed and long-term pharmacovigilance agreed. Phase 4 post-marketing studies have become pivotal to the monitoring of DMTs outwith clinical trials and offer the only reasonable method of identifying unanticipated long-term safety issues. In TRANSFORMS (comparing Fingolimod with Avonex®), for example, there were 2 cases of breast cancer and 3 cases of melanoma in the 0.5mg Fingolimod group and none in the Avonex® group. The authors conclude that the numbers are too small to draw definitive conclusions about causality, hence raising the concern of RCTs being underpowered to demonstrate differences between cohorts when rare, but serious, events are considered. In this study, there were also 3 cases of Basal Cell Carcinoma (BCC) in the Fingolimod cohort versus 1 in the interferon group: since Fingolimod has been licensed, it has become apparent that BCC occurs at a higher rate with Fingolimod use than would be expected and, hence should be actively monitored for in treated patients. Similarly, the development of PML in patients using Fingolimod and Dimethyl Fumarate, after their licensing, in patients never treated with Natalizumab provides further evidence of the need for post-marketing safety monitoring and should be borne in mind by clinicians discussing newer DMT options[128,129]. The apparent parallel increase in efficacy and safety issues is noted and some suggest that risk of adverse events is a proxy for efficacy, for the currently licensed treatments at least. As understanding of the cause(s) of MS improves and treatments become more targeted, it could be anticipated that unexpected adverse events become less prevalent.

Treatment failure definitions

A model for escalation of treatment to more efficacious, higher risk treatments is only rational if treatment failure can be defined. However, there remains

uncertainty as to what constitutes a failure of treatment in terms of efficacy and, therefore, when escalation of treatment should be considered[130,131]. There are a number of studies which have aimed to determine prognostic factors based on early clinico-radiological outcomes in treatment trials and, hence, identify which patients are likely to benefit from a change in treatment.

The decision to change DMT is based on the balance of tolerability and efficacy. Intolerable side effects may lead to a 'horizontal' switch to a DMT of similar efficacy and safety profile; lack of efficacy would suggest the need to introduce a higher efficacy treatment ('vertical' switch), albeit with a potentially less favourable safety profile. It is the 'vertical' switch to higher efficacy DMTs that is a particular subject under review here, but 'horizontal' switching to an agent with a different mechanism of action has a biological basis and is employed by some clinicians.

The difficulty in quantifying the number of relapses or new MRI lesions 'allowed' before escalating treatment (in other words, defining 'treatment failure'), is borne of the fact that no DMT is completely successful in suppressing such measures of disease activity, and RRMS symptoms are episodic by definition. In a given individual the assumption that that individual would have more lesions/relapses/disability were they not on treatment may be entirely incorrect. Further difficulties arise from the time course and individual nature of the history during DMT treatment – for example when a DMT is well tolerated and perceived to have been relatively efficacious prior to a change in disease status (such as a patient relapsing after 4 years of relapse-free therapy), whether to switch, and what to switch to, can present particular challenges.

In clinical practice a number of clinical and radiological factors govern the decision about whether to escalate to higher efficacy treatments – these 'markers of disease activity' are imperfect and limited, but are the best guide available in routine practice currently. A number of groups have attempted to provide specific guidance on the level of worsening that suggests treatment failure and a need to switch, based on post-hoc analysis clinical trial data [132,133134], observational studies[135,136] or expert opinion[137]. All agree that development of new relapses and MRI lesions are markers of treatment failure, but there is no clear consensus on how to quantify these. The Canadian group[137]

also include disability progression (as measured by EDSS) in their decision-making model, but this necessitates detailed objective examinations at two time-points, which is not always undertaken in routine clinical practice.

The Rio score[138] was based on an observational cohort of 222 RRMS patients and used 1-year data to predict outcome at 2 years, including MRI, relapse and disability criteria. The Modified Rio Score (MRS) was adapted from this to include only relapses and MRI, using 4-year data from a pivotal trial of IFNß-1a[139]. MRI findings after 1 year of treatment are dichotomised into either ≤4 new T2 lesions or more. Scores of 0-2 were given for 0, 1 or 2 relapses respectively and, combined with the MRI score, had prognostic value for disability progression within the next 3 years. Hence, it was argued that >4 new T2 lesions and 1 or more relapses after 1 year of treatment would suggest the need for an alternative strategy where possible, as these patients could be considered 'non-responders'.

Bermel et al.[132] assessed patients from another pivotal placebo-controlled IFN trial[140], 15 years after their participation in the 2-year randomised phase, in a bid to identify predictors of long-term outcome for those in the treatment arm. Disease activity whilst on treatment, as measured by MRI and clinical relapses during the 2 year-trial, predicted the severity of disability 15 years after completion, despite the fact that treatments differed among patients after the trial. These data suggested that >3 new T2 lesions, ≥2 clinical relapses and ≥2 Gd-enhancing lesions during the 2 –year treatment phase predicted incrementally higher odds of severe disability after 15 years. This reinforced the importance of disease control in the early stages, as well as suggesting potential parameters to guide treatment response (or lack of it).

Using data from the Avonex-Steroids-Azathioprine (ASA) Study, completed in 2003, Horakova et al. sought to identify responders and non-responders to IFN treatment based on the relationship between year 1 events and outcomes 2-6 years later[134]. Of the original 181 patients included in the trial, 172 had complete datasets for this post-hoc analysis. All patients had been treated with Avonex® but were subdivided such that 56 (32.5%) had no other treatment, 55 (32%) took Avonex® in combination with azathioprine (50mg od) and 61 (35%) took Avonex® combined with azathioprine and prednisolone (10mg alternate

days). Since completion of the original study, 126 (73%) had treatment changes either during the 1st year of the study (n=3) or in the intervening 6 years (n=123). The majority discontinued treatment it seems, with others having treatment added or switching to an alternative first-line therapy, but the raw numbers quoted in the paper do not tally with the total cohort who changed. Median time on any DMT during the 6-year follow-up was 6 years (IQR 62-72 months) i.e. the population was largely treated for the majority of follow-up analysed. Notably, the study was ultimately negative in that the addition of azathioprine or prednisolone was not associated with different efficacy in comparison to IFN alone[141] – therefore, this population can be considered comparable to IFN-only treated patients it seems. The authors retrospectively divided the follow-up period into a 'Prediction Phase' (Year 1 of the study) and a 'Response Phase' (Years 2-6 thereafter) in order to identify early predictive factors, in treated patients, which result in poorer outcomes i.e. non-responders. An Annualised Relapse Score was calculated based on the number and severity (based on impact on ADLs) of relapses occurring during follow-up. Patients were initially defined as non-responders if EDSS increased from baseline, sustained until the entire duration of follow-up, and the Annualised Relapse score was >1 on average during follow-up occurred. A 2-stepped EDSS increase was used for this definition i.e. increased by >1 if baseline EDSS ≥1 or by1.5 if baseline EDSS=0. Thirty-six patients (21%) met both criteria for non-response, but the analysis is based on 90 patients (52%) who only fulfilled one of the pre-defined criteria. From these patients, the number of new MRI T2 lesions and Annualised Relapse Score within the first year were predictive of non-response to IFN treatment. EDSS change was not a statistically significant predictor of non-response. Relapse score and the number of T2 lesions were further analysed to provide odds ratio for each new T2 lesion or 1-point increase in Relapse Score in the first year. Greater than or equal to 1 new T2 lesion during the first year of IFN was associated with a 3.3 times higher risk of being a non-responder [OR 3.3 (95% CIs 1.6-6.6)] in comparison to no new lesions. Interestingly, ≥3 new T2 lesions was associated with a lower, but still increased, risk of being defined as a non-responder [OR 2.8 (1.4-5.5)]. Greater than or equal to 1 relapse was associated with a higher risk of non-response [OR 2.7 (1.5-5.0)] and this increased with the number of relapses [OR 3.4 (1.4-8) for ≥3 relapses]. Combined MRI and relapse activity was strongly predictive of being a

non-responder: a relapse score of ≥2 *and* ≥2 new T2 lesions provided an Odds Ratio of 18 (5-67)]. In terms of MRI changes alone, patients with ≥3 new T2 lesions within the first year of treatment were at higher risk of non-response (OR=3) regardless of relapse activity. This study therefore provides evidence that combined relapse and MRI activity within the first year of IFN treatment predicts future relapses and sustained disability in the medium-term. Additionally, it provides useful numerical cut-offs, similarly to Rio and Bermel, for the number of MRI lesions which may be tolerable in the first year of treatment. Notably, patients with relapse activity but no new T2 MRI lesions in the first year were not at increased risk of future relapses or disability in this study. Of course, it may be the case that these clinico-radiological events predict disability in the long-term, which is not assessed by this analysis. The authors also note that, in keeping with other studies, a lower EDSS at baseline (≤2) was associated with better outcomes overall. Again, this suggests that treatment earlier in the course of disease (or at least before disability is established) is likely to be beneficial.

In a prospective observational study based at a single Italian centre Prosperini et al. also found that MRI changes within the first year of IFN treatment were predictive of later disability but the magnitude differed[135]. The aim was to compare MRI criteria alone with the guidance issued by the European Medicines Agency (EMA), which requires at least one relapse and ≥9 T2 lesions or ≥1 contrast-enhancing lesion in order to pursue escalation to 2nd-lne treatments. A total of 370 patients were followed-up for 4 years from an initial cohort of 445. Patients who received IFN for <1year (n=65, 14.6%) were excluded: 38 (8.5%) discontinued treatment for lack of efficacy. All available IFN preparations were included and analysed as a single cohort. In this analysis, the number of relapses (not including severity) and the 2-step EDSS worsening used by Horakova et al. were the primary outcome measures upon which treatment failure was based. Overall, the EMA guidelines and MRI changes within the first year both similarly predicted higher relapse rate and disability worsening in comparison to patients without disease activity. One or more new contrast-enhancing lesion or ≥2 T2 lesions were more sensitive [61% (95%CIs 54-67) vs. 34% (27-41)] and accurate [70% (65-74) vs. 57% (51-62)] than the EMA guidelines, but both were similarly associated with a higher risk of relapse and disability worsening in the following

4 years (EMA guidelines: HR=3.69 for relapses, HR=6.02 for disability worsening; MRI criteria: HR 3.15 and 5.31 respectively). These data suggest that MRI activity alone is a reasonable proxy for the EMA criteria and, hence, escalating in the absence of relapses, when there is ≥1 new contrast-enhancing or ≥2 new T2 lesions within the first year of IFN treatment, could be considered. Notably, the specificity increases with increasing numbers of new T2 lesions [specificity 96% (92-98) for ≥3 new T2 lesions] but sensitivity is reduced [31% (24-37)], meaning a higher number of new T2 lesions is unlikely in those who will have better outcomes, but does not definitely predict a worse outcome in those in which this occurs. Similarly, the combined predictive value of relapses and MRI activity improves specificity but not sensitivity. It can therefore be fairly reassuring that those without MRI or relapse activity in the first year of IFN will have relatively better outcomes over the next 4 years but that, whilst more likely, it does not necessarily follow those with clinico-radiological activity are destined to disability. MRI disease activity should, however, certainly be considered when escalation decisions are being made.

The largest, and most recent, dataset employed to define treatment failure using IFN comes from the multicentre European MRI in MS (MAGNIMS) network[136]. This included 1280 patients with RRMS from 9 European centres with prospectively collected data but retrospective analysis. Again, clinico-radiological events during the first year of treatment with any IFN were analysed for their association with treatment failure. This study included just 2 years of follow-up, however. Treatment failure was also defined as EDSS worsening, confirmed at 6 or 12 months, again using a 2-step paradigm based on the baseline EDSS as in previous studies. Additionally, treatment was considered to have failed if the patient had been escalated to a second-line therapy (because of lack of efficacy, not tolerability) within the following 2 years, at the discretion of the treating physician and their patient. Time-to-treatment-failure was then used as the outcome variable for a multivariate Cox model where the number of relapses and new T2 lesions in the first year were assessed for their relative contribution to the outcome. This then allowed estimation of the average effect of the number of new T2 lesions and relapses during the first year on the risk of treatment failure in the subsequent 2 years. The authors found that there was significant heterogeneity between centres with regard to T2

lesions resulting in a treatment failure but that this did not appear to invalidate their results. However, one of the centres was excluded as a result of its different approach to the others in this respect. Unfortunately, the proportion of patients who were deemed treatment failures due to EDSS worsening and (physician-directed) treatment escalation are not provided. The factors upon which treatment escalation were based were not standardised, or reported, therefore the inclusion of this as a definition of treatment failure suggests a circular analysis. However, analysis of EDSS worsening as an outcome measure mirrors that of treatment failure suggesting it is robust. Patients were grouped into 3 tiers of increasing disease activity, from no relapses and <3 new T2 lesions in the first year to 1 relapse and ≥3 new T2 lesions *or* ≥2 relapses. In this way, it was demonstrated that those with the most clinical and/or radiological disease activity had significantly shortened time to treatment failure and EDSS worsening. Additionally, multivariate Cox modelling permitted Hazard Ratio calculations for the effect of each incremental T2 lesion or relapse in the first year on the risk of treatment failure at 3 years. Patients with no new T2 lesions or relapses in the first year respectively were used as the reference point for comparison. The number of new T2 lesions at which the increased risk of treatment failure at year 3 became significant was 4 [HR 2.36 (95%CIs 1.35-4.16)]. A single relapse was associated with a HR of 1.84 (1.39-2.44) for treatment failure at 3 years and ≥2 relapses 3.03 (2.06-4.45) in comparison to no relapses. Again, this reiterates the predictive value of MRI in isolation but the occurrence of early relapses is also predictive of more future disease activity and that the combination of both is even more likely to predict treatment failure as defined here.

The Canadian Treatment Optimisation model[137] grades disease activity based on relapses, MRI activity and disability, and defines 3 levels of 'concern'. As well as including disability measures, this guidance also categorises relapse severity, rather than the absolute number alone, and the year of treatment in which they occur. A single attack in the 2nd year of treatment with minimal effect on ADL, no motor or cerebellar involvement and prompt recovery without the need for steroids would suggest a 'low' level of concern with regards to disease activity; more than 1 attack in the first year of treatment, with severe motor involvement and incomplete recovery at 6 months would be of 'high' concern. Similarly, EDSS changes are graded and the Timed 25-Foot Walk (T25FW) is included, in contrast to the other scoring systems. A single new Gd-enhancing or

T2 lesion (per year) on MRI is considered 'low' concern but ≥3 of either suggest a 'high' concern. A cumulative level of concern is then reached based on these findings and the higher the level of concern, the stronger the rationale to escalate therapy.

Overall, then, the definition of treatment failure is largely based on evidence from IFN use and suggests that early relapses and MRI activity beget short- and medium-term poorer disease outcomes, generally suggestive of the need to escalate treatment. The cut-offs differ slightly but most studies suggest any relapse activity within the first year of treatment and >3-4 new T2 lesions (even in the absence of relapse activity) or >1 new contrast-enhancing lesion are associated with more disease activity in the following 2-6 years, in terms of relapses and disability worsening. The desire to define non-responders early (i.e. within the first year of treatment) results in a more limited role for EDSS change given the small proportion of patients having a significant change over this relatively short period – but its presence is rightly concerning and may inform escalation decisions, particularly if thought to be part of an inflammatory phenotype.

A practical point which is not considered in these studies is the time to effectiveness of the agents involved. Pivotal DMT trials generally show a separation of the treatment and control arms at 3-6 months, in terms of clinical efficacy outcomes, hence it is generally agreed that clinical or radiological disease activity within the first 3-6 months of treatment does not necessarily reflect a lack of efficacy of the treatment but, rather, it has simply not reached full efficacy in this time. Of course, it is the case that the purported time to efficacy of DMTs simply reflects the fact that clinical and radiological assessments first occur at these timepoints, rather than this being a pharmacodynamic fundamental. Indeed, it has been postulated that Copaxone® specifically has a longer latency to effectiveness (up to 9 months) but this is challenged by a recent pharmaceutical-sponsored post-hoc analysis of a GA placebo-controlled trial, which suggests treatment efficacy is evident even at 2 months[142]. Similarly, patients may mount an antibody-response to treatment which, if transient, delays onset of full efficacy but does not appear to have an ongoing impact unless the neutralising antibodies are persistent[143].

'Real-world' studies are the primary (Class IV) evidence of various escalation options available with the increasing development of DMTs in the past decade. Prosperini et al. evaluated 285 patients with RRMS from 2 Italian centres who had failed treatment with IFN or Copaxone and were escalated to Natalizumab or switched to another IFN/Copaxone®[144]. Treatment failure was defined as ≥2 relapses or 1 relapse with residual disability after at least 1 year of treatment. Although the study was prospective, over 2 years, patients were not randomly assigned to each treatment cohort: 'the choice of second-line treatment depended on the availability of Natalizumab or was a patient choice due to safety concerns regarding PML'. After patients and their physicians had chosen whether to pursue escalation or a switch to an alternative 1st-line DMT, a propensity score-adjusted Cox regression model was used to balance the groups. There were significant differences between the unmatched groups, the escalation cohort having higher relapse rates and EDSS at baseline. Additionally, a higher proportion of the escalation group had failed on high-dose IFN or Copaxone® whereas the switch group had more patients failing on low-dose IFN. These factors suggest the group which ultimately escalated had worse disease, as would be expected. Most of the switch group (62.6%) actually simply increased their dose of IFN, rather than changing to a new DMT, however. In total, 106 (93%) of patients in the escalation group and 161 (94%) in the switch group completed 2 years of follow-up. In the non-adjusted analysis, the escalation group had significantly higher proportions of patients with absence of relapse activity, EDSS progression and MRI activity at 24 months (but there was no difference at 12 months). This could reflect regression to the mean, given their higher baseline disease activity, but the statistically matched patients also showed lower hazard ratios for disease activity in the escalation group, including No Evidence of Disease Activity (NEDA, defined as no relapses, progression or MRI inflammatory lesions) in comparison to those who switched between immunomodulators [HR 0.51 (95%CIs 0.35-0.74), p=0.001].

Using the MSBase and Tysabri Observational Program (TOP) datasets, Spelman et al. also reported improved outcomes for patients escalating to Natalizumab, after relapse on 1st-line DMTs, in comparison to those who did not[145]. MSBase is a large prospectively collected dataset from 73 countries, largely European, currently with almost 50000 patients included. The TOP dataset includes

longitudinal follow-up of Natalizumab-treated patients, mainly for pharmacovigilance. Interrogating these databases, the authors identified patients who had relapsed on first-line treatments ('BRACE') in the 12 months prior to switching to either an alternative BRACE treatment or Natalizumab. Over 4500 patients available within these databases met inclusion criteria: propensity-matching was undertaken to improve comparability between the cohorts, including indication bias, resulting in 869 matched pairs of patients. Covariates used to generate propensity scores and, hence, cohorts comparable on known confounders, included gender, age, disease duration, baseline EDSS, number of DMT initiations, duration of DMT use (as a proportion of disease duration) and relapse rates in the past 12-24 months at baseline. The matching process was successful at balancing known confounders between the cohorts. Mean follow-up was longer (2.24 years) in the BRACE group than in those escalating to Natalizumab (1.95 years), but duration between follow-up visits was comparable. Overall, annualised relapse rate was significantly lower for 4 years in patients switched from BRACE treatments to Natalizumab rather than another BRACE treatment (0.14 vs 0.36, p=0.0002) but the difference was greatest and most sustained in those switching from IFN than with Copaxone (or both). After year 1, for example, there was no significant difference in those who switched to Natalizumab from Copaxone®, in terms of ARR but the numbers available for these sub-group analyses are notably smaller. Disability outcomes were based on 374 patients continuing BRACE treatments after relapse and 514 escalated to Natalizumab with 3 EDSS scores available. Patients who escalated to Natalizumab had a 26% reduction in the risk of 3-month confirmed disability progression [HR 0.74 (95%CIs 0.55-0.97, p=0.036] but there was no significant difference in time to disability progression within the first 12 months. Using the area under the curve method, to capture EDSS changes over time, the escalated group had significantly lower levels of disability over 24 months of follow-up overall, however. Additionally, patients escalating to Natalizumab were significantly less likely to discontinue treatment in comparison to those remaining on BRACE treatments [HR 0.26 (0.2-0.34), p<0.001]. This study did not evaluate safety outcomes but it is well established that PML risk is significantly higher with Natalizumab use and has not been reported with any of the BRACE treatments. These data therefore, again, support the conclusion that escalating to Natalizumab, after disease activity on first-line treatments, is superior in

terms of efficacy outcomes than switching amongst first-line treatments but with higher associated risk.

The benefit of escalation has not been demonstrated in all studies however. In a recent single-centre Italian study, no significant difference was found in the proportion of patients reaching NEDA at 24 months, whether they switched to another first-line DMT or a second-line DMT after clinical disease activity on first-line therapies[146]. This retrospective analysis of their local database identified patients with RRMS who had a relapse and/or confirmed disability progression within 12 months preceding a switch from IFN/Copaxone to Copaxone/IFN or a second-line DMT (Fingolimod, Natalizumab, Cyclophosphamide) and with subsequent 12 months of follow-up available. Treatment use was based on prescriptions dispensed but drug use was not confirmed or measured over time. The study included 91 patients, 48 of who performed a 'lateral' (horizontal) switch and 43 escalated. There were significant baseline differences in MRI lesion load between the groups, the escalated group having more. Otherwise the groups were comparable but no attempts were made to match them further. The primary endpoint was the proportion reaching NEDA (no relapses, progression or MRI activity) at 24 months and this was not significantly different between the groups (20.8% in switch group vs. 18.6% in escalation group). Similarly, time to EDSS 4 and first relapse was comparable between the groups. The numbers of patients within the 24-month analysis small (10 and 8 in escalation and switch NEDA groups respectively) and clearly the escalation group had higher MRI disease burden, signalling a worse prognosis and preventing statistically-meaningful comparisons between the groups.

Similarly, other data from MSBase suggests that switching to purported higher efficacy oral DMTs from first-line treatments does not improve efficacy outcomes when the switch is made for reasons of tolerability, rather than efficacy[147]. This retrospective study used propensity-score matching to identify 396 comparable pairs of patients (N=792) with stable RRMS for 12 months who switched to either an alternative injectable or oral DMT largely due to side effects or inconvenience from an initial IFN/Copaxone. The cohorts were well matched using propensity scores and most switched to Fingolimod (71.2%) over DMF (16.2%) or Teriflunomide (12.6%) in the oral switch group. Despite this,

there was no difference within the first 6-months between the groups in terms of relapse rates or disability progression whether patients switched to an oral or alternative injectable DMT. Additionally, there was no significant difference between relapse rates between the three available oral agents when adjusted for baseline differences. Concluding that switching to oral DMTs in the setting of poor tolerability of injectable treatments does not improve efficacy is mitigated by the short follow-up in this study and the acknowledgement that the reason for switching was not documented in over half of the patients included.

Overall then, the benefits of an induction approach are favourable against the risks when used at an early stage of disease when there is significant disease activity and patients may wish to accept higher risks for the purported benefits. An escalation approach remains the prevailing paradigm in MS therapeutics and the first-line injectable agents have a long-standing proven safety record, albeit with a higher risk of MS disease activity. Treatment failure should be considered when any relapse activity occurs within the first year of treatment or >3-4 new T2 lesions (even in the absence of relapse activity) or >1 new contrast-enhancing lesion as these are associated with more disease activity in the following 2-6 years in terms of relapses and disability worsening. Clinical and MRI activity within the first 3-6 months of treatment may simply reflect subtherapeutic pharmacodynamics of the agent used and hence caution should be exercised in considering this a treatment failure. The available evidence suggests that escalating to a second-line DMT is more effective than switching between first-line DMTs if treatment failure occurs, but there is no evidence of an efficacy benefit in switching to oral DMTs from injectable DMTs for tolerability reasons alone. In choosing the escalation agent after first-line treatment failure, Alemtuzumab is more effective than Rebif® and Natalizumab appears more efficacious than Fingolimod but individual patient risk-factors, such as co-morbidities and JCV status, should of course be considered carefully when such choices are being made.

1st- and 2nd-line DMTs in Clinical Practice: Real-world data

Dimethyl Fumarate and Fingolimod

The merit of randomised controlled trials (RCTs) is obviously well-founded, but the artificial nature of cohort selection, concordance and follow-up reflects the efficacy of an intervention under ideal conditions and hence outcomes cannot necessarily be extrapolated directly to real-world practice[148]. Similarly, the chronic nature of MS requiring decades of follow-up to demonstrate long-term benefits or safety issues does not lend itself to a RCT model given both the practical and ethical issues around blinding, as well as the financial costs and inevitable attrition of participants over such long periods. That said, the RCT is undoubtedly the least biased approach given its ability to balance both known and unknown confounders such that any difference observed between cohorts can be attributed to treatment effects with greater confidence. There is clearly a role for both methods in reaching conclusions about the safety and efficacy of DMTs in RRMS as long as the limitations of both are understood. Considering these limitations in the design of treatment studies in MS, as well as analysis, can minimise their impact or at least rationalise any conclusions. With this in mind, we sought to describe the efficacy and safety of DMF when used in our centre and specifically considering post-marketing safety concerns.

Dimethyl Fumarate (DMF) is an immunomodulatory treatment for RRMS which demonstrated efficacy in two placebo-controlled randomised trials[64,65]. Its therapeutic effect is thought to be related to activation of the nuclear factor (erythroid-derived2)-like (Nrf2) pathway, with beneficial modulation of pro- and anti-inflammatory cytokines. It was the second licensed oral treatment for RRMS and the first available for use as a first-line treatment. After its approval in April 2014, our centre became one of the leading users of DMF in Europe.

Fingolimod is a first-in-class, sphingosine-1 phosphate (S1P) receptor modulator, and was the first oral disease modifying treatment (DMT) licensed for use in multiple sclerosis (MS)[149]. Fingolimod alters trafficking of both naïve and antigen-activated lymphocytes from secondary lymphoid tissue, thymus and bone marrow to produce a relative lymphopaenia. It does this via two non-specific mechanisms - an agonistic action on S1P receptors 1,3,4 and 5, and a functional antagonistic action against the SIP1 receptor. The cumulative effect

of these actions is the internalisation of the S1P1 receptor within lymphocytes, thereby resulting in the transient retention in lymph nodes. The non-selective mode of action of fingolimod may produce unwanted additional effects, for example bradycardia. Combined analysis of FREEDOMS and TRANSFORMS showed a mean decrease in heart rate by 8 beats per minute, and an incidence of 1st degree AV block of 4.7%. This is thought to occur as a consequence of fingolimod's transient agonistic effects on S1P1 receptors within atrial myocytes[150]. Originally evaluated as a treatment for renal transplant rejection, at doses 10 times higher than now prescribed in MS, in renal transplant trials macular oedema was noted to be twice as prevalent in patients receiving fingolimod compared to those on placebo (28% in diabetic patients and 4% in non-diabetic patients). All subsequent clinical trials involving fingolimod have therefore implemented screening for macular oedema[151]. The incidence of macular oedema from the TRANSFORMS and FREEDOMs trials was 0.3% for 0.5mg Fingolimod, and 1.2% for the 1.25mg dose, and the vast majority of cases occurred within 3-4 months of starting the study drug[66,123]. The manufacturer of fingolimod (Gilenya®) has recommended ophthalmological evaluation is undertaken at 3-4 months after treatment initiation[152]. Further, it is recommended that multiple sclerosis patients with diabetes mellitus or a history of uveitis undergo an ophthalmological evaluation prior to initiating therapy and have follow-up evaluations while receiving therapy. It is recommended that Gilenya be discontinued if a patient develops macular oedema. Fingolimod has been available for use in the UK since 2012. In Scotland, it is licensed for use as a first-line DMT in patients with highly active relapsing remitting MS (RRMS) or failing a 1st-line therapy due to ongoing disease activity. [153]

As outlined previously, clinical trials do not provide efficacy and safety results which are necessarily comparable to real-world outcomes once the drug is used in a wider population, and this was the case for both DMF and Fingolimod with regards to progressive multifocal leukoencephalopathy (PML)[154,155]. PML had been associated with Fumaderm, the parent drug of DMF used mainly for psoriasis in Europe, but no cases occurred during the DMF or Fingolimod clinical trials for RRMS. The importance of post-marketing data was illustrated by the subsequent diagnosis of PML in a patient receiving DMF with notably prolonged lymphopaenia and further case reports have since also suggested a link between

prolonged lymphopaenia and risk of PML with DMF[154]. Lymphopaenia occurred in approximately 37% of patients in the pivotal trials of DMF, around 7% having Grade 3 [ALC (Absolute Lymphocyte Count) <0.5x10^9/l] at least once during follow-up[156]. As a result of the association between lymphopaenia and PML with DMF from these post-trial events, European guidelines now recommend considering DMF discontinuation if grade III lymphopaenia is present for 6 months. On stopping DMF, however, lymphopaenia can persist and global data on the expected recovery of ALC following DMF discontinuation due to lymphopenia is limited to a small number of patients from pivotal trials showing that, whilst lymphocyte counts generally increase on discontinuation, levels had not returned to pre-treatment levels after 4 weeks. Risk factors for Fingolimod-associated PML have not been defined.

With these issues in mind, we undertook firstly an analysis of the overall utility (real world effectiveness) of DMF and Fingolimod in our centre. Additionally, we wished to evaluate real world safety outcomes relevant to each product, specifically lymphopaenia and lymphocyte recovery with DMF since this is the chief identifiable PML risk factor, and has major implications for the introduction of subsequent therapy. Speed and level of lymphocyte recovery (in any population) was unknown from trial data available at the time, and so the analysis of this particular safety aspect represented data particularly pertinent to contemporary practice. We also wished to evaluate rates of macular oedema with Fingolimod in a real-world population, rather than the artificial trial populations studies in the pivotal studies – again vital data for practising clinicians.

Methods

Retrospective analyses of electronic medical records were undertaken in our regional tertiary Neurology centre. The Institute of Neurological Sciences in Glasgow is the neurosciences referral centre for the West of Scotland, covering a catchment area of 2.5 million people. Each NHS patient in Scotland is assigned a unique identifier, the Community Health Index (CHI) number, which is included in admissions records and hence permits access to their electronic medical records including GP referrals, laboratory results and all secondary care correspondence. Additionally, the electronic portals are linked in the West of Scotland such that regional records are available centrally and vice versa,

permitting access to all follow-up correspondence even for patents located out with our tertiary centre. A contemporary electronic prescription, from GP records, is also accessible to view current and recent medications which patients have been prescribed.

Patients treated with DMF were identified through local prescription records. Firstly, all patients who commenced DMF during its first year of availability were identified and, secondly, all patients who had discontinued DMF due to lymphopaenia between December 2014 and March 2016.

All patients commencing Fingolimod are admitted to our day investigation unit for first-dose cardiac monitoring. Patients were identified using local prescription and admission records. In the first analysis, all patients treated with fingolimod in our centre by April 2015 were identified and their clinical records used for retrospective analysis of tolerability and efficacy. Later, we identified all patients commenced on fingolimod in our centre between May 2016 and May 2017 specifically evaluating screening for, and occurrence of, fingolimod-associated macular oedema. Patients admitted for re-initiation of fingolimod after a break in treatment were not included in the evaluation of macular oedema.

Demographic, clinical, radiological and laboratory data were collected and analysed using Microsoft Excel® and XLSTAT®. Descriptive statistics were compiled. Parametric data were compared using a 2-tailed t-test, non-parametric data with the Mann-Whitney test and proportions were compared using z-scores.

This work includes data collection and analyses completed by Frederick Winslow (University of Glasgow), Kieran Fitzpatrick (Pharmacist, NHS GGC) and Sarah-Jane Martin (Neurology Registrar, NHS GGC).

Results

Dimethyl Fumarate

A total of 156 patients (74% female) commenced DMF within its first year of availability in our centre. **Table 1-1** summarises the demographic, clinical, radiological and laboratory results for the cohort. There were no significant differences in radiological or

laboratory outcomes between those starting DMF first-line or having had previous treatment(s). As expected, those starting DMF first-line had a shorter disease duration in comparison to those having had other previous treatments but time to first relapse on treatment was significantly shorter for treatment-naïve patients.

Lymphopaenia and discontinuation due to side effects were higher than in the pivotal clinical trials. In our cohort, 46% (66/143) developed a lymphocyte count less than 0.8 $x10^9$/l on at least one occasion during the first year of treatment. Additionally, 20.5% (32/156) of our cohort discontinued DMF due to side effects. No renal impairment, proteinuria or opportunistic infections were documented in our cohort.

Table 1-1: DMF cohort characteristics and outcomes

	Whole Cohort (N=156)	DMF 1st Treatment (N=36)	DMF not 1st Treatment (N=120)
n (%) Female	116 (74)	25 (69)	91 (76)
Mean Age (SD) [years]	39.5 (8.9)	38.3 (11.8)	39.8 (7.9)
Mean time since diagnosis (SD) (years)	5.6 (6.2) (n=131)	**3.5** (8.0) (n=36)	**6.5*** (4.9) (n=95)
Number of DMTs tried before DMF [n (%)]		1	72 (60)
		2	31 (26)
		3	17 (14)
RELAPSES			
Relapse occurred [n (%)]	21 (13.5)	8 (22.2)	13 (10.8)
Disabling relapse occurred [n(%)]	7 (4.5)	3 (8.3)	4 (3.3)
Steroids used [n(%)]	5 (3.2)	1 (2.8)	4 (3.3)
Mean duration to 1st relapse after starting DMF (SD) [months]	6.3 (3.7)	**4.4** (3.7)	**7.5*** (3.3)
MRI FINDINGS			
MRI undertaken [n(%)]	54 (34.6)	16 (44.4)	38 (31.7)
Reason for MRI			
Re-baseline [n(%)]	24 (44)	4 (25)	20 (52.6)
Interval [n(%)]	14 (26)	4 (25)	9 (23.7)
Following Relapse [n(%)]	16 (30)	8 (50)	8 (21.1)
Mean duration since starting DMF (SD) [months]	8.4 (3.5)	8.3 (4.1)	8.5 (3.6)
Median time since comparison MRI (IQR) [years]	1.4 (2.0)	1.3 (0.8)	1.7 (3.2)
Patients with new T2 lesions [n(%)]	23 (42.6)	8 (50)	15 (39.5)
Re-baseline [n(%)]	10 (41.6)	0	8 (40)
Interval [n(%)]	3 (21.4)	2 (50)	3 (33.3)
Following Relapse [n(%)]	10 (62.5)	6 (75)	4 (50)
Patients with new Gd-enhancing lesions [n(%)]	7 (13)	4 (25)	3 (7.9)
LYMPHOCYTE NADIR			
FBC checked at least once [n(%)]	143 (91.7)	34 (94.4)	109 (90.8)
No lymphopaenia [n(%)]	32 (22.4)	5 (13.9)	25 (22.9)
1.1-1.5 $x10^9/l$ [n(%)]	45 (31.5)	13 (36.1)	32 (29.4)
0.8-1.0 $x10^9/l$ [n(%)]	29 (20.3)	6 (16.7)	23 (21.1)
0.5-0.7 $x10^9/l$ [n(%)]	31 (21.7)	5 (13.9)	26 (23.9)
<0.5 $x10^9/l$ [n(%)]	6 (4.2)	3 (8.3)	3 (2.8)
CONTINUED ON TREATMENT AT TIME OF ANALYSIS			
n (%)	109 (69.9)	23 (63.9)	86 (71.7)
Mean duration of DMF treatment (SD) [months]	14.1 (1.5)	13.9 (1.4)	14.2 (1.5)

DISCONTINUED TREATMENT BY TIME OF ANALYSIS			
n (%)	47 (30.1)	13 (36.1)	34 (28.3)
Mean duration of DMF treatment (SD) [months]	6.2 (4.2)	8.4 (3.8)	5.3 (4.1)
Reason for stopping			
Side effects [n(%)]	32 (20.5)	7 (19.4)	25 (20.8)
Lymphopaenia [n(%)]	9 (5.8)	2 (5.6)	7 (5.8)
Treatment failure [n(%)]	4 (2.6)	3 (8.3)	1 (0.8%)

*p<0.05

The majority of patients (75.8%) commenced DMF having failed at least one previous DMT. **Table 1-2** shows the number of patients using each DMT and the sequence used. The majority (n=72, 60%) had been initiated onto an injectable treatment before being switched to DMF. Indeed, where DMTs were sequenced before switching to DMF, the injectable therapies were mostly used. A smaller proportion of patients switched from 2nd-line therapies (Fingolimod / Natalizumab), largely having failed other DMTs first. **Table 1-3** outlines the reasons for switching to DMF from the various DMTs. Side effects were the main reason to switch from all treatments except Natalizumab, where, of the 6 patients switching from this, half were due to treatment failure and half due to PML concerns because of JCV antibody positivity. Additionally, a significant proportion of patients using Rebif switched because of concerns regarding thrombotic microangiopathy (TMA) which had occurred in patients in Scotland around this time [157] and all patients were counselled on the risks and offered the option to switch.

Table 1-2: DMTs used prior to DMF (N=120)

	1 DMT tried (n=72)*	2 DMTs tried (n=31)		3 DMTs tried (n=17)		
		1st DMT**	2nd DMT	1st DMT	2nd DMT	3rd DMT
Avonex®	24 (34.2%)	12 (40%)	4 (12.9%)	5 (29.4%)	5 (29.4%)	4 (23.5%)
Betaferon®	0	0	0	1 (5.9%)	1 (5.9%)	0
Extavia®	0	0	0	0	1 (5.9%)	0
Rebif®	23 (32.9%)	10 (33.3%)	7 (22.6%)	5 (29.4%)	5 (29.4%)	2 (11.8%)
Copaxone®	23 (32.9%)	7 (23.3%)	12 (38.7%)	5 (29.4%)	5 (29.4%)	5 (29.5%)
Fingolimod	0	0	4 (12.9%)	0	0	4 (23.5%)
Natalizumab	0	1 (3.3%)	4 (12.9%)	1 (5.9%)	0	2 (11.8%)

*2 unknown

** 1 unknown

Table 1-3: Reasons for switching to DMF from *last* DMT tried

DMT	Reason for switching (to DMF) [n (%)]					Total
	Injection issues	Side effects	Preferred oral treatment	Treatment failure	Other	
Avonex						
1st DMT	5	10	3	6	0	**24 (75%)**
2nd DMT	0	3	0	1	0	**4 (12.5%)**
3rd DMT	1	0	0	1	2	**4 (12.5%)**
Total	**6 (18.8%)**	**13 (40.6%)**	**3 (9.4%)**	**8 (25%)**	**2 (6.3%)**	**32**
Rebif						
1st DMT	7	5	4	2	5	**23 (71.9%)**
2nd DMT	1	0	1	0	5*	**7 (21.9%)**
3rd DMT	0	1	0	0	1	**2 (6.3%)**
Total	**8 (25%)**	**6 (18.8%)**	**5 (15.6%)**	**2 (6.3%)**	**11 (34.4%)**	**32**
Copaxone						
1st DMT	12	4	2	5	0	**23 (57.5%)**
2nd DMT	5	2	2	3	0	**12 (30%)**
3rd DMT	2	1	0	0	2	**5 (12.5%)**
Total	**19 (47.5%)**	**7 (17.5%)**	**4 (10%)**	**8 (20%)**	**2 (5%)**	**40**
Natalizumab						
2nd DMT	N/A	0	0	1	3**	**4 (66.7%)**

3rd DMT	N/A	0	0	2	0	**2 (33.3%)**
Total	**N/A**	**0**	**0**	**3 (50%)**	**3 (50%)**	**6**
Fingolimod						
2nd DMT	N/A	4	N/A	0	0	**4 (50%)**
3rd DMT	N/A	2	N/A	1	1	**4 (50%)**
Total	**N/A**	**6 (75%)**	**N/A**	**1 (12.5%)**	**1 (12.5%)**	**8**
Overall Total	33 (28.0%)	32 (27.1%)	12 (10.2%)	22 (18.6%)	19 (16.1%)	**118¶**

*Concerned about TMA

**JCV positive after > 2yrs Rx

¶Data missing for 2 patients

'Other' reasons included patient concern about TMA with Rebif (n=4), JCV positivity and lack of response to 2 other DMTs (n=1), Natalizumab antibodies positive (n=1) and unknown reason/unattainable from available clinical documents (n=5).

The second analysis identified all patients who discontinued DMF due to lymphopaenia between December 2014 and March 2016. At this time, 594 patients had been commenced on DMF in our centre. A total of 30 patients (5.1%) discontinued DMF due to lymphopaenia. Median age for those discontinuing due to lymphopaenia was 44.5 years (range: 31-60), in comparison to 39.5 for the entire cohort. Similarly, there was a higher proportion of females in the discontinuation group compared with the overall cohort (80 vs 74%). DMF was the first DMT for almost half (n=14) but was second- or third-line in 37% (n=11) and 16% (n=5) respectively. Median ALC before DMF use was $1.5x10^9/l$ (range 1-3.1) and $0.6x10^9/l$ (range 0.2-0.8) at the time of stopping, with 73% (n=22) JC positive. DMF was given for a median of 331 days (Range 169-597). Discontinuation occurred with Grades I, II and III lymphopaenia in 2(6%), 23(78%) and 5(16%) patients respectively.

Median follow-up after DMF discontinuation was 3 months (range 0-14). The majority of patients (18, 60%) commenced a new DMT. Patients given fingolimod

(n=8) and interferons (n=8) after DMF had a subsequent decline in ALC. Six patients (20% of those discontinuing and 1% of the entire cohort) had prolonged lymphopaenia after DMF discontinuation, defined as ALC ≤0.9x10^9/l for ≥6 months, despite no further treatments. Notably, time to lymphopaenia development after DMF commencement was shorter in those who subsequently developed prolonged lymphopaenia after its discontinuation in comparison to patients who did not (median 4 vs. 7 months).

Therefore, female sex, older age and shorter latency to lymphopaenia after DMF initiation appeared associated with prolonged lymphopaenia after DMF discontinuation. ALC prior to starting DMF, number of treatments prior to DMF and time on DMF was not. Additionally, in our centre, JC Virus antibody positive status appears a factor in the decision to discontinue DMF when lymphopaenia occurs during treatment.

Fingolimod

Safety

93 patients had been commenced on fingolimod in our centre by April 2015. The mean duration of fingolimod therapy was 415 days (SD 261). All of the patients in our cohort used fingolimod as a 2nd-line DMT, whereas this was the case for less than half in FREEDOMS (42.6%) and TRANSFORMS (44.8%). Our cohort is compared with those in the pivotal fingolimod trials in **Table 1-4**. Older and with a higher proportion of female patients, our cohort also had a shorter disease duration and most had used one or more DMT before fingolimod, perhaps suggesting a more active disease phenotype than the trial cohorts. In our cohort, 59 patients (63%) switched from injectable therapies due to disease activity on treatment and 23 (25%) switched from natalizumab due to JCV Ab positivity and, hence, PML risk with continued natalizumab treatment.

Table 1-4: Fingolimod cohort characteristics

Baseline Characteristics		**Study Cohort (n=93)**	**TRANSFORMS Cohort (n=431)**	**FREEDOMS Cohort (n=425)**
Age (SD) [Years]		38.3 (9.5)	36.7 (8.8)	36.6 (8.8)
Female (%)		69 (74)	282 (65)	296 (70)
Time since diagnosis of MS (SD) [Years]		6.7 (5.1) [n=84]	7.5 (6.2)	8.0 (6.6)
Number of DMTs pre-fingolimod (% cohort)	1	55 (59)	-	
	2	31 (33)	-	
	3	7 (8)	-	

Nineteen patients (20.4%) discontinued fingolimod in our cohort, with the majority (14/19, 15.1%) due to side effects. Side effects leading to discontinuation were headache (4/14), lymphopaenia (3/14), diarrhoea (2/14) and 1 patient each for back pain, limb pain, rash. Additionally, 1 patient had bradycardia during initiation therefore fingolimod was not continued and 1 patient was found to have macular oedema during follow-up screening hence fingolimod was discontinued. In comparison, 80/425 (19%) discontinued fingolimod in total, but with just 35 (8%) due to an adverse event or abnormal laboratory results in FREEDOMS and 6% in TRANSFORMS.

The overall occurrence of side effects in our cohort are outlined in the **Table 1-5**. Again, headache was the most common and comparable to the frequency seen in the pivotal trials.

Table 1-5: All documented side effects related to Fingolimod

Side effect	Study cohort (n=93)	%	FREEDOMS Population (n=425)	%	TRANSFORMS Population (n=429)	%
Headache	25	26.9	107	25.2	99	23.1
Infections	22	23.7	-	-	-	-
UTI	10	10.8	34	8.0	26	6.1
LRTI	6	6.4	41	9.6	-	-
URTI	3	3.2				
Herpes virus	2	2.2	37	8.7	9	2.1
Dental	1	1.1	-	-	-	-
Vaginal Thrush	1	1.1	-	-	-	-
Raised GGT	21	22.6	67*	15.8	-	-
Depressed mood	12	12.9	33	7.8	21	4.9
Palpitations	9	9.7	-	-	-	-
Diarrhoea	8	8.6	50	11.8	32	7.5
Reported worsening vision	7	7.5	-	-	-	-
Transient chest Pain	4	4.3	4	0.9	-	-
Dizziness	4	4.3	31	7.3	24	5.6
Dyspepsia	4	4.3	38	8.9	40	9.3
Itch	3	3.2	-	-	-	-
Constipation	3	3.2	-	-	-	-
Back pain	3	3.2	50	11.8	26	6.1
Breakthrough bleeding	2	2.2	-	-	-	-
Reduced appetite	2	2.2	-	-	-	-
Rash	2	2.2	-	-	-	-
Macular oedema**	1	1.1	0	0	-	-
Bradycardia**	1	1.1	9	2.1	2	0.5
Tachycardia**	1	1.1	-	-	-	-
Aching limbs**	1	1.1	-	-	21	4.9

***FREEDOMS reports ANY "abnormal liver function tests"**

**** Side effect reported by a single patient only reported if leading to discontinuation of therapy or admission/medical review of fingolimod treatment**

Fingolimod-associated macular oedema

In the first analysis of those starting fingolimod before April 2015, only 1 patient developed macular oedema as outlined above. Subsequent further analysis of patients who commenced on fingolimod in our centre between May 2016 and May 2017 was undertaken to investigate the ongoing evaluation and occurrence of fingolimod-associated macular oedema (FAME) in more detail. This identified 40 patients who had initiated onto fingolimod in this time. The mean age for this cohort was 39 years (SD 10.9) with a range between 19 and 60 years. The majority were female (32, 80%). No patients were on medications predisposing to macular oedema but 1 patient (2.5%) had diabetes. Duration of available follow-up correspondence was at least 6 months from fingolimod initiation, with a mean treatment duration of 351.8 days (SD 98.2).

All patients were referred and seen by ophthalmology to screen for macular oedema. The patient with diabetes was referred and appropriately screened prior to commencement of fingolimod. All patients underwent OCT as part of the ophthalmological assessment.

Mean time from fingolimod initiation to ophthalmology assessment was 124.8 days (SD 46.8), range 28-238]. Two patients did not attend their first appointed ophthalmology assessment: excluding these patients (duration from fingolimod initiation to ophthalmology assessment 182 and 199 days respectively) reduced the mean duration of the cohort to 121.4 days (SD 45.4). Overall, however, 25 patients (62.5%) waited more than the advised 120 days between fingolimod initiation and ophthalmology screening for macular oedema.

In considering reasons for delays, the mean duration between fingolimod initiation and ophthalmology referral was 15.6 days (SD 15.8) but up to 50 days in one case. Two patients were ‘pre-referred’ prior to fingolimod dosing notably, although one still waited beyond 120 days. Generally, however, the greatest delay was awaiting the ophthalmology appointment, with mean 109.3 days (SD 41.6) from referral, but up to 230 days in one case and 17 patients (42.5%) waiting more than 120 days for this alone, as outlined in **Figure 1-1**.

Macular oedema was found in 2 patients (5%), neither of whom had a predisposing condition or drug therapy. These patients were older than the

cohort average (44 and 53 years), with one male and one female, but the limited numbers preclude any inferences about this. Both had switched from a previous DMT (Glatiramer Acetate and Dimethyl Fumarate) but both screened after the target 120 days (129 and 136 days respectively). Neither patient had visual symptoms. In one case, there was minimal intraretinal fluid and no discernible visual impact hence this was managed expectantly with continuation of Fingolimod and had resolved spontaneously during follow-up over 2 months. There have been no visual sequelae documented since but, interestingly, the patient was diagnosed with Type 2 diabetes within 6 months after this finding. The second patient had significant unilateral macular oedema and measurable acuity reduction (6/12+2) despite a lack of visual symptoms. Fingolimod was stopped immediately and the OCT changes had resolved within 3 months.

The majority of patients were switching to fingolimod from natalizumab, due to PML risk with the latter, but fingolimod was being used first-line in a significant proportion as outlined in **Figure 1-2**. This suggests a trend toward using fingolimod first-line, which was not the case in the earlier cohort.

Figure 1-1: Duration between fingolimod initiation and ophthalmology review for FAME

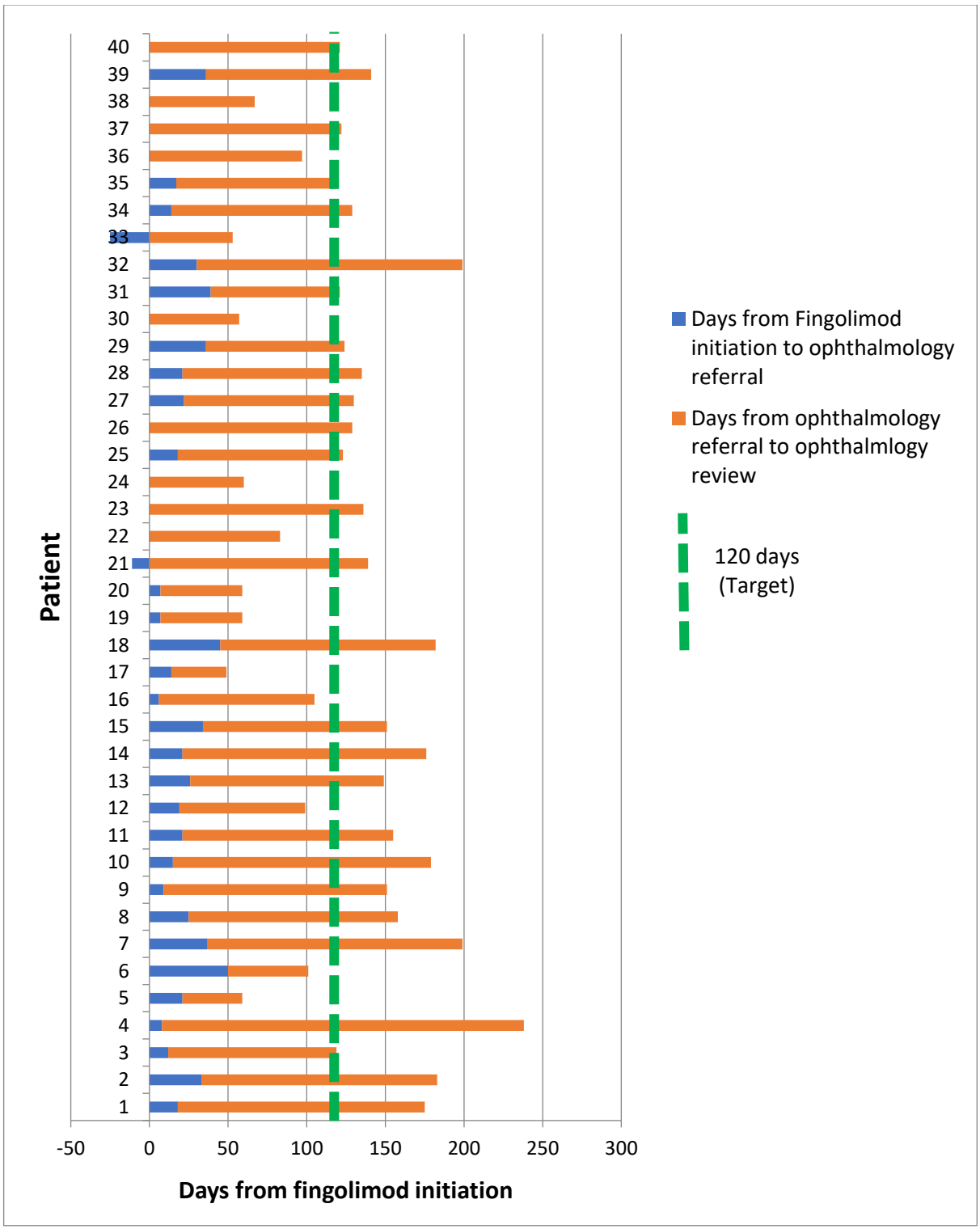
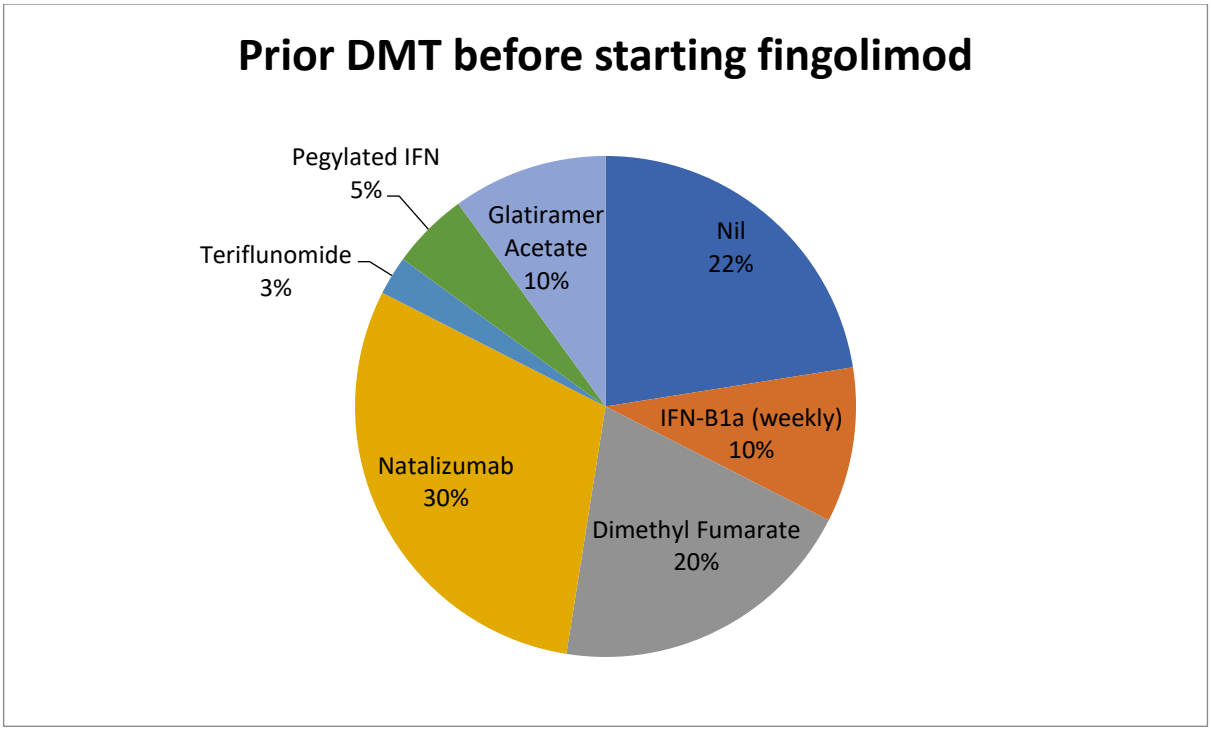


Figure 1-2: DMT use prior to fingolimod (N=40)



Efficacy

Efficacy of fingolimod in real-world practice was assessed using basic outcomes for the 93 patients started on fingolimod by April 2015 in our centre. The cohort was split into 3 groups:

- A: switchers from BIFN/GA (n=56)
- B: switchers from natalizumab (n=24)
- C: other circumstances (for example, those starting fingolimod after treatment interruption)

In terms of clinical efficacy, 73 patients (78.5%) remained relapse-free overall, after a mean treatment duration of 407 days (SD 257), and this was similar whether switching from an injectable DMT or natalizumab. This was within the range seen in TRANSFORMS (82.5% after mean 360 days) and FREEDOMS (70.4% at mean 720 days). The ARR for all fingolimod patients was 0.27 (SD 0.74, range 0.12-0.42), which did not differ significantly whether switching from injectables or

natalizumab but was higher than that reported in the pivotal trials and may reflect a more active cohort from baseline or simply the real-world use of treatment where concordance and monitoring are less vigorous than in a clinical trial setting. **Table 1-6** tabulates these results.

Table 1-6: Relapse activity with fingolimod use in our centre vs pivotal trials

	Study Cohort			TRANSFORMS (N=429)	FREEDOMS (N=425)
	Switching from injectable (n=56)	Switching from Nat (n=24)	Total (N=93)		
Relapse-free (%)	44 (78.6)	19 (79.2)	73 (78.5)	354 (82.6)	299 (70.4)
ARR (SD)	0.22 (0.55)	0.39 (1.1)	0.27 (0.74)	0.16 (0.42)	0.18 (0.31)

Of the 93 patients studied, 86 had a pre-fingolimod MRI, mean 279 days (95% CIs 192.7-365.4) prior to drug initiation. The first subsequent follow-up MRI was available for 63 patients and occurred after a mean of 329 days (95% CIs 267.9-389.6). The majority of patients (38/63, 58.7) had no evidence of either increased T2 lesion load or contrast-enhancing lesions. Of those who did develop radiological disease activity in this time, 22 (34.9%) had new T2 lesions and 4 (6.4%) had new contrast-enhancement. This is comparable with the findings from TRANSFORMS, where 54.8% of patients had no increased T2 lesion load or new contrast-enhancing lesions, 35.3% developed new T2 lesions and 9.9% contrast-enhancement over a similar time period.

Discussion

The available literature pertaining to DMT treatment in RRMS suggests benefits in early initiation and escalation with real-world observational studies complementing randomised controlled trials. Data from our centre has shown some differences in safety and efficacy outcomes with DMF and fingolimod in

comparison to published studies, as well as identifying possible predictors of prolonged lymphopaenia with DMF.

Clearly, there is a spectrum of disease at presentation, from RIS to PPMS, and the ideal time to start treatment to maximise benefits and minimise risks remains debated, with resultant variations in practice. A particular issue is identifying an appropriate outcome measure to demonstrate benefit, particularly with short-term trials in a lifelong disease, and many agree that the past focus on inflammatory activity and motor-heavy disability measures such as the EDSS is unsatisfactory. There are sound immunopathological reasons to suggest benefit of early treatment and MRI modalities demonstrate the loss of brain tissue at even the earliest stages, suggesting a target for end-organ damage even at this stage. Clinical trials have shown clear discrepancy in the benefits of treatment in early (RR) and late (SP) MS but at what point in an individual's disease course that anti-inflammatory treatment becomes futile remains uncertain and is dependent on clear definitions of disease stages and agreement on the most valid outcome measures, which remain debated. Indeed, global variance in advice from national regulatory bodies on the use of DMTs in CIS for example, based on the same evidence, belies the uncertainty within the MS community regarding their utility. Despite purported long-term benefits from early DMT initiation, much of the evidence is subject to methodological flaws and the role of real-world studies to enrich the evidence-base is becoming more accepted but have their own limitations too. The reality is that there are a number of factors influencing the timing and choice of DMTs, including safety, cost, patient factors, clinician factors and the regulatory environment in which the decisions are made.

It is clear, however, that DMTs are not fully effective or always tolerated so changing treatments is often necessary. The decision to 'escalate' to a different DMT presupposes that some therapies are more effective although there is a paucity of head-to-head randomised trials comparing many DMTs and even fewer evaluating escalation directly in those failing treatment. This approach remains the most accepted presently, but early aggressive 'induction' therapy is being more widely used, particularly in those with highly active disease, and offer the possibility of avoiding long-term maintenance treatments and the need to switch between them. Again, definitions are important and there remains no single

agreed definition on the magnitude of disease activity that should trigger the offer of treatment change, but data suggest those who have relapses or new MRI lesions early after starting treatment are likely to benefit from a change. The introduction of PROMs will help establish whether there are true benefits perceived by patients with such approaches, and DMT use in general, and it seems likely that a multimodal outcome measure of MS will be most appropriate whilst the specific aetiology of MS remains elusive.

Real-world evidence from observational studies clearly provide insight into disease and its management that randomised controlled trials (RCTs) cannot always provide[44,148]. Although RCTs can establish the efficacy of an intervention under ideal conditions, they do not necessarily provide the best indication of its effectiveness in real-world practice. As explained in the earlier chapter, Montalban describes this as the difference between *efficacy*, treatment outcomes under the ideal conditions in a clinical trial setting, and *effectiveness*, the outcomes actually achieved in everyday practice. Patients included in RCTs are not necessarily representative of real-world populations and their limited follow-up duration make them unsuitable for long-term efficacy and safety outcomes. Real-world studies include subgroups not included in RCTs and can provide long-term data not otherwise available. Additionally, particularly in multicentre observational studies, large populations can be included and comparisons between various DMTs can be made. However, real-world studies are subject to numerous biases but strategies exist to minimise these as far as possible, in both the planning of the study and the statistical analysis. Indeed, real-world studies have already provided useful results predictors of treatment response[158], comparative effectiveness[159] of DMTs and long-term effectiveness[11].

With these issues in mind, we evaluated the use of DMF (N=156) and Fingolimod (N=93) in patients with RRMS in our centre. Both treatments were typically used after the failure of injectable therapies, largely due to side effects for DMF but breakthrough disease for fingolimod. In the DMF cohort, time to first relapse on treatment was significantly shorter in those using DMF 1st-line in comparison to those switching from another therapy (4.4 vs 7.5 months) but there was no significant difference in efficacy or safety outcomes between these subgroups otherwise. The occurrence of lymphopaenia and treatment discontinuation due to side effects in our cohort was higher than that reported in the pivotal trials of

DMF. Additionally, we found that increased age, female sex and a shorter latency to lymphopaenia onset were associated with prolonged lymphopaenia after DMF discontinuation but the sample size is too small to be definitive and confirmation of this would be needed in a larger cohort. In our fingolimod cohort, safety and efficacy outcomes were largely comparable to that in its pivotal trials, but we identified some adverse events which hadn't been reported albeit in small numbers. Additionally, similarly to DMF, the proportion of patients discontinuing fingolimod due to adverse events was higher than in the trials. ARR was also higher in our real-world cohort than in fingolimod trials but this may relate to baseline differences between the cohorts, as much as any impact from uncontrolled real-world factors. In a specific evaluation of fingolimod-associated macular oedema (N=40) screening, 2 patients (5%) were identified during recommended screening which is higher than the 0.7% in FREEDOMS and 0.5% in TRANSFORMS but clearly these had a much larger sample size. We also identified delays to screening for a majority of patients which may be amenable to local remedial measures with time.

Our local data on DMF and fingolimod use confirm the utility of real-world observational data and some of the limitations of randomised trials as outlined above. The higher rates of treatment discontinuation due to side effects in our real-world cohort than pivotal trials may reflect the different motivation or acceptance of patients and clinicians in a clinical trial versus an everyday setting where the levels of monitoring, follow-up and time investment differ, for example. The identification of rare adverse events will only be identified by large real-world, longer term follow-up studies but are extremely important in identifying the true benefits and costs of a treatment. Our identification of patients potentially at higher risk of prolonged lymphopaenia after DMF discontinuation needs further confirmation but could act as a springboard for other centres to evaluate this also and speaks to the goal of personalised medicine such that the decision to embark upon DMF may be evaluated differently in the older female patient who may need further DMTs after DMF, were these findings to be replicated on a larger scale. Lastly, the evaluation of local factors in implementing the use of a new treatment e.g. fingolimod-associated macular oedema is crucial to ensure the appropriate framework and

infrastructure are in place in order to use treatments safely and effectively in the real world.

Conclusions

The current literature on DMTs in RRMS support the utility of their use early in the disease course for maximum benefits and switching to alternatives when clinical or radiological disease activity occurs. A number of questions remain as to how early in disease DMTs should be introduced, what treatment goals should be, what constitutes treatment failure and the comparative efficacy of available DMTs, particularly in the long-term. The limitation of randomised clinical trials to answer these, and other clinically-relevant questions, is recognised and the role of observational real-world studies is expanding, as well as understanding of their own limitations and strategies to minimise these.

We describe outcomes for patients who started DMF during its first year of availability in our centre (N=156) and all patients treated by 2016 (N=594). The majority of patients commenced DMF due to tolerability issues with injectable treatments. The proportion of patients discontinuing DMF due to side effects and experiencing lymphopenia was higher than that reported in pivotal trials. There was a significantly shorter latency to first relapse on DMF for treatment-naïve patients in comparison to those switching from a previous therapy. The occurrence of lymphopaenia with DMF was evaluated in the larger cohort after it had been used in our centre for a few years (N=594). Of these patients, 5.1% discontinued DMF due to lymphopaenia, the majority of these being JCV Ab positive. Older female patients were over-represented in the lymphopaenia group and 6 patients (20% of the discontinuation cohort, 1% of the total treatment group) had prolonged lymphopaenia for 6 months after DMF discontinuation despite no further treatments. Shorter latency to lymphopaenia after starting DMF was associated with subsequent prolonged lymphopaenia after DMF discontinuation. This has potential implications for instituting new DMTs after DMF discontinuation, in terms of PML risk for example, hence prolonged lymphopenia and its attendant risks should be kept in mind even after DMF is stopped.

Two cohorts of fingolimod-treated patients were also evaluated – all patients commenced on treatment by April 2015 (N=93) and those starting treatment between 2016 and 2017 (N=40). Efficacy and safety outcomes were evaluated in the former group and the occurrence and screening process for macular oedema in the latter. Similarly to DMF, a higher proportion of patients discontinued fingolimod due to adverse events than expected from trial data but that efficacy results were largely reassuring, although our cohort had generally failed injectable therapies (which was not the case for trial patients) and may belie more active disease and explain some discrepancies. Patients in our centre often waited longer than advised for macular screening after fingolimod initiation and, although a higher incidence was found than in the trials, there was no deleterious effects where macular oedema was identified. Local practice can be reviewed in order to minimise delays in ophthalmological assessment and overall these data provide a comprehensive picture of the local experience to help inform patients and clinicians as to the reality of DMT use.

In the end, different study methodologies should be used to complement each other and provide a more holistic view of the benefits and costs of MS treatments which patients and clinicians can then use to inform decision-making in practice. The next chapter considers the approach to the use of DMTs in regional MS centres in Scotland and resultant outcomes.

Chapter 2: Multiple sclerosis Outcome Determination Evaluating Real Differences After TimE (MODERATE)

Introduction

Uncertainty regarding the efficacy and safety of DMTs for RRMS, as well as their cost, has led to varied practice in their use by clinicians since their introduction in the early 1990s. Few would suggest that DMTs have no place in the management of RRMS, but the indications for initiation and switching between therapies, and timing of their use in the course of the disease, remains contentious. Large randomised controlled short-term clinical trials have established a reduction in relapse frequency and short-term disease worsening using a variety of DMTs. Long-term extension studies suggest ongoing effects on disease activity, but are flawed by a number of methodological difficulties and cannot properly address the impact of treatment initiation and early escalation during the initial phase of the disease on longer-term impairments, which are often the main cost to patients and society. The extent to which early treatment affects such fixed impairments remains the fundamental issue in MS therapeutics.

The debate about balancing efficacy with safety has intensified with the advent of higher efficacy DMTs and their attendant safety concerns, and the identification of significant but rare safety concerns with lower efficacy treatments such as cases of PML (progressive multifocal leukoencephalopathy) in patients taking dimethyl fumarate (DMF). Improved imaging and the establishment of clinical and radiological prognostic factors early in the disease course that correlate with long-term disability help guide clinicians in making difficult treatment choices, but the best choice of therapy in the individual remains inexact and uncertain. Long-term randomised controlled trials are both impractical and arguably unethical in the current climate, and real uncertainty exists regarding whether early treatment policies affect the outcomes that really matter to people with MS.

Within Scotland, different centres and specialists have different approaches to the use of DMTs. The purpose of this study is to compare outcomes between patients cared for by these clinical teams, who differ only in respect of the clinician overseeing their care. We aim to address one of the fundamental unanswered questions in MS care: to what extent does a policy of early inflammatory disease control influence real clinical outcomes? Notably, the

purpose of concentrating on data from the Grampian and Glasgow centres is to enrich the sample with prognostically similar but differently managed patients, rather than comparing the centres.

The Scottish MS Register (SMSR) was used to identify patients for this study. The SMSR is a national Register within the Scottish Healthcare Audits programme at the Information Services Division (ISD) of NHS National Services Scotland (NSS). The aim of the SMSR is to improve healthcare for people living with MS in Scotland. The SMSR includes all patients newly diagnosed with MS in Scotland since 1st January 2010. The inclusion criterion is a new diagnosis of MS, after 1st January 2010, as defined by Poser (1983) and the Revised McDonald Criteria (2005). Patients with 'possible' MS or a clinically isolated syndrome are not included. Data is held in accordance with NSS data protection guidelines and includes demographic information, date of first symptoms, family history of MS, diagnostic categorisation, investigations undertaken and MS nurse involvement after diagnosis.

The most recent SMSR report at the time our study began (September 2015), includes 2164 patients diagnosed with MS in Scotland since 2010. Case ascertainment is thought to be very high, with around 90% of the true number of patients diagnosed with MS in Scotland registered. All patients who make contact with MS services are included, minimising referral bias. Notably, whilst the register itself does not include all clinical data potentially relevant to the patient or proposed research questions, all NHS patients in Scotland have a unique identifier, the Community Health Index (CHI) number. The SMSR includes the CHI number, which is used in all community and hospital health records in Scotland, thus opening the potential for a much broader scope of relevant data access and collection, both clinical and radiological, for our study.

There is uncertainty as to the true utility of DMTs in MS, specifically their long-term benefits, and what constitutes treatment success and failure[160-163]. The primary outcome measure in pivotal clinical trials is almost uniformly the (relative) reduction of relapse frequency in the comparator arm, over a relatively short period of time (2-3 years), in a lifelong disease.

The rationale for relapse reduction as a treatment goal in RRMS is derived from large natural history studies which suggest predictable, irreversible progression of disability once a particular level of disability is reached[31,164]. It is proposed that prevention of relapses, and therefore prevention of the accrual of disability that occurs as a consequence of each relapse, will prevent or delay the development of irreversible fixed disability. The level of disability is described by the Expanded Disability Status Scale (EDSS) score, ranging from 0 (asymptomatic and normal neurological examination) to 10 (death due to MS). The scale is non-linear and lower scores may result from examination findings alone, in the absence of reported disability, whilst higher scores reflect more apparent fixed disability, principally the need for walking aids and mobility limitation. Particular 'milestones' within this scale are recognised: while the time to deteriorate from a score of 0 to 3 is highly variable between individuals, once an EDSS score of 3 is reached predictable irreversible progression of disability occurs[17]. For this reason, we have chosen an EDSS score of 3 as our primary efficacy endpoint to evaluate the utility of early DMT initiation and escalation.

Whether the reduction of relapses and inflammatory activity prevent the occurrence of irreversible disability milestones and irreversible progression is uncertain. While many experts share the view that early intervention with DMTs will lead to longer term benefits, the 'causal' relationship between relapse rate and disability progression is easily challenged[165] along with DMT trial designs in general[43,45], leaving clinicians uncertain how to apply trial data in clinical practice.

The degree to which DMTs alter likelihood and severity of permanent disease accumulation and progression in MS remains the most clinically important question in modern MS practice, and is unlikely to be answered by traditional randomised studies, given the impracticalities of designing trials that would have to last decades. The typical timescale of a clinical trial does not provide an opportunity to assess long-term fixed disability reliably. Reduction in 3- or 6-month sustained increases in disability scores during 2- year trials are claimed to represent long-term positive effects of DMTs from clinical trial data, but their validity is highly questionable[166].

Conflicting evidence suggesting no effect of DMTs (IFN-ß) on disability progression[167] and a significant effect on disease progression[168], with similar clinical endpoints, has been published. Recent evidence from the DMT risk-sharing scheme in the UK[11] provides some 'real world' data as to the merits of DMTs in the long-term, but the outcome measures used and the validity of the control group lead to real uncertainty and caution regarding a firm conclusion. The published data have thus not answered the central question facing clinicians dealing with early stage RRMS, and there remains a climate of uncertainty regarding the true long-term benefit of early intervention, and whether this is outweighed by the inevitable risk of immunomodulatory therapy

More recently, the concept of a 'window of opportunity' for MS treatment has emerged following experience with newer DMTs, as discussed earlier, which clearly impacts DMT initiation strategies. Trials of the anti-CD52 monoclonal antibody Campath (now known as Alemtuzumab) failed to show any significant effect on disability progression in patients with established disability in the progressive phase of their condition, despite reducing relapses and radiological markers of inflammation; the opposite was the case for those with early active MS, where a significant reduction in the accumulation of disability was observed in treated patients[99]. Additionally, clinical and imaging studies now support longstanding pathological data suggesting that the neurodegenerative aspect of multiple sclerosis (as measured by brain atrophy and cognitive dysfunction) occurs in parallel with clinico-radiological inflammatory disease[169,170] at the onset of the disease, is predictive of long-term disability[171,96] and may be ameliorated by DMT use[172,173,174,175]. These factors inform the practice of many clinicians in advocating the early initiation of DMTs in the disease course; indeed, the use of the highest efficacy agents at the earliest stage, 'induction' (as discussed earlier), follows intuitively from such observations. The highest efficacy DMTs (in terms of relapse reduction) are associated with higher risk in terms of short- and long-term adverse events. In addition, the limitations in our ability to predict disease course on an individual basis, given the heterogeneity of MS, result in understandable caution on the part of both clinician and patient when considering this strategy.

The limitations in our ability to predict an individual's disease course from the onset of RRMS has a major impact on both the patient's and clinician's

willingness to accept risk from early initiation of DMTs. The heterogeneity of prognosis for patients with RRMS is well established, with patients having widely different levels of disability despite the same duration of disease. Irreversible limitation in ambulation, a unilateral aid for walking and becoming wheelchair-bound occurred at median times of 8, 20 and 30 years respectively when examined in the era before widespread use of DMTs[13], but such summary statistics do not provide a clear prognosis to the individual. It is becoming accepted that the concept of 'benign' MS is either a misnomer or rarer than previously described[176,177], the definition having relied largely on motor dysfunction and not accounting for other disabling symptoms suffered by MS patients in the long-term, particularly cognition[178,179]. In instituting DMTs, particularly those with significant risk of adverse events, an accurate individual prognosis would allow thoughtful patient selection to identify those most likely to benefit. As outlined previously, a number of predictive prognostic factors have been identified through natural history studies[13,180,181] and all agree that the nature of the disease in its earliest stages has profound effects on the development of irreversible disability even decades later. A high number of relapses in the early years, particularly with motor or sphincter involvement, incomplete recovery after the first relapse, early accumulation of disability and a short interval between the first and second relapse are recognised as predictive of a higher risk of long-term disability. Age, gender, sensory symptoms and ataxia at onset are less clearly relevant in predicting long-term outcomes [181].

More recently, statistical methods have been employed with the aim of developing useful clinical prognostic models to characterise a patient's likely disease course as early as possible[182]. The BREMSO (Bayesian Risk Estimate for Multiple Sclerosis at Onset) score is a composite of some of the prognostic factors described above and has been shown to predict MS outcomes successfully. Importantly, the BREMSO score has good specificity but poor sensitivity – patients with a low BREMSO score are unlikely to develop severe disease but high scoring patients may well have a favourable disease course. The BREMSO scoring system appears to work best at the extremes, but is less useful in predicting those with intermediate risk, which will of course be the majority.

The timing of DMT initiation is controversial and practice varies around the world. In the United States and Canada, DMTs are typically commenced after a clinically isolated syndrome (CIS) of demyelination in the absence of dissemination in time to meet the current diagnostic definition for MS[183]. In Europe, and certainly the UK, DMTs would not usually be initiated until dissemination in time and space has been confirmed clinically, after a second relapse, or radiologically on MRI. The ability to make a diagnosis of MS after a single event and a single initial MRI is a relatively recent one, since the updated McDonald criteria in 2010[183].

Proponents of DMT initiation after CIS cite the high likelihood of conversion to CDMS (the onset of which has been shown to be delayed by DMTs in numerous large RCTs), that early treatment of RRMS has been proven to be beneficial, and that the 1st line DMTs have an excellent safety profile [184]. Conversely, the potential of exposing those who would otherwise not have a second event or whose disease natural history would have been 'benign' to DMTs, the lack of proven benefit of DMTs in the longer term, the modest benefits in disability outcomes with early treatment (which reduce with time) and the significant financial costs of DMTs are reasons that support a 'watch and wait' approach[185]. The decision to institute DMTs once the MS diagnosis has been made is standard practice in many westernised nations, but the timing of treatment, and the effects that early treatment has on real outcome, remain contentious.

The choice of whether to begin treatment with first- or second-line DMTs is largely dictated by the guidance issued by national advisory bodies, following an appraisal of the clinical benefits and financial costs of different treatments, which determines their 'cost-effectiveness', in different clinical groups. In Scotland, the Scottish Medicines Consortium (SMC) advises health boards as to the cost effectiveness of DMTs and the clinical indications for which they can be used. SMC advice takes account of the available evidence and applies to Greater Glasgow and Grampian, so providing an equal framework for decision-making in both regions. Despite this, great differences exist between different clinicians and different departments in their use of DMTs, both in Scotland and in the UK as a whole – the guidelines are not prescriptive, do not mandate treatment in particular scenarios, and allow for very significant differences in interpretation (for example, what constitutes 'a disabling' relapse, what constitutes 'a

significant' increase in scan lesion load). This means that individual clinicians have considerable autonomy in prescribing practice, hence real behaviour in the prescription of DMTs is dictated by each clinician's interpretation of the conflicting and incomplete data on treated and untreated prognosis as described above.

Whilst attempts to predict the efficacy of DMTs have been the subject of numerous randomised controlled studies, predicting their safety in individual patients has largely been influenced by initially unforeseen post-marketing safety concerns. An example of this has been the development of Progressive Multifocal Leukoencephalopathy (PML) in patients treated with Natalizumab for RRMS[186,187] , which has led to a global surveillance and risk stratification programme, based on the JC virus (JCV) exposure of the patient, the duration of treatment and previous use of immunosuppressants[188]. The unanticipated occurrence of PML and other serious adverse events informs the decision-making process of clinicians and patients in the use and choice of DMTs at the outset and throughout the disease course. This is particularly the case for higher efficacy and newer treatments with less robust long-term safety profiles, and is a factor pertinent to the choice and timing of DMT initiation and escalation in the cohorts studied in MODERATE.

Notably, the options available for DMT initiation changed during the period under study. Fingolimod, Teriflunomide, DMF and Alemtuzumab were only approved in Scotland by the SMC as first-line therapies in 2014. Prior to this, Fingolimod could only be used if a patient had failed on IFN or GA, and the other agents were not licensed. Thus, at the time we are assessing DMT initiation in our study (2010-11), Natalizumab was the only high efficacy treatment licensed and approved for use first-line (in patients with particular disease characteristics), and the vast majority of DMTs initiated were either interferons or GA.

Methods

There were 2 parts to this study:

1) Identifying whether patients with similar disease (based on PS) are treated differently in Scotland in terms of DMT initiation

2) Establishing whether the difference in treatment impacts on clinico-radiological outcomes

Aim

To provide evidence for clinicians and patients regarding current practice in Scotland and the benefits and risks of early initiation of disease modifying treatments (DMTs) in relapsing-remitting multiple sclerosis (RRMS), using real world data.

Primary Objective

Establish whether patients with similar severity of RRMS are treated differently within Scotland and establish a cohort of differently treated patients with similar severity of RRMS, for comparison of prospective outcomes after diagnosis.

Primary Endpoint

1) Mean DMT initiation propensity scores in patients initiated onto a DMT within 1 year of diagnosis and those not initiated onto a DMT within 1 year of diagnosis

Secondary Objective

Compare *retrospective* safety and efficacy outcomes in cohorts generated by the calculation of propensity scores

Secondary Endpoints

Propensity-score matched pairs were categorised into 'treated' and 'not treated' groups and retrospective comparisons of the following were made by review of medical records:

1) Proportion of patients reaching EDSS 3 first 5-6 years after diagnosis in each group

2) Proportion of patients relapsed in first 5-6 years after diagnosis in each group
3) Annualised relapse rate in first 5-6 years after diagnosis in each group
4) Proportion experiencing (serious) adverse events in first 5-6 years after diagnosis in each group

Cohort identification, and sample size projections

This was a retrospective observational study using routinely collected clinical data. Our original intention had been to identify differently treated patients with RRMS, despite similar disease prognosis at diagnosis, and invite them for a detailed clinico-radiological assessment to compare outcomes over time between those treated and untreated. Unfortunately, funding for the prospective phase of the study was not obtained until 2019, and so the methods and results presented here pertain only to the completed (retrospective) part of the study.

Prior to data collection, we hypothesised the various patient trajectories from diagnosis to treatment in order to define the groups we would compare for clinicoradiological outcomes and estimate the likely sample sizes. This was done using available SMSR data and review of the relevant literature to estimate treatment and disease characteristics. Between 2010 and 2011, there were 292 patients registered on the SMSR with a new diagnosis of MS from the Greater Glasgow and Grampian regions. In Greater Glasgow, there were 98 new diagnoses in 2010 and 91 in 2011; in Grampian, there were 54 and 49 diagnoses of MS respectively. Around 90% of all MS diagnoses follow a relapsing-remitting course[4,189] (RRMS), so we estimated a total sample of around 263 patients with RRMS from this cohort, which would form our total sample (n=170 from Greater Glasgow; n= 93 from Grampian, N=263).

Available data from a national patient survey[190] suggested that patients with similar demographic, clinical and radiological characteristics will have been managed differently in the two centres. The average of the reported value for DMT use in Grampian from this recent survey[1] (19.4%) suggested that, for Grampian patients diagnosed with RRMS in 2010-11 (n=93), 75 (80.6%) may be eligible but not taking DMTs; in Greater Glasgow the survey untreated rate is

54%, hence we estimated 92 of the 170 patients diagnosed in 2010-11 would be untreated, giving a total of 167 (75 + 92) initially untreated patients. An average of 46% of eligible patients were taking DMTs in Greater Glasgow according to the survey meaning, of our initial cohort of 170, 78 patients were estimated to be taking DMTs in Glasgow, with 19% (18) of the patients in Grampian, giving a total of 96 treated patients. **Figure 2-1** below illustrates these calculations. For the purposes of these calculations, we assumed those starting treatment do so within 1 year of diagnosis.

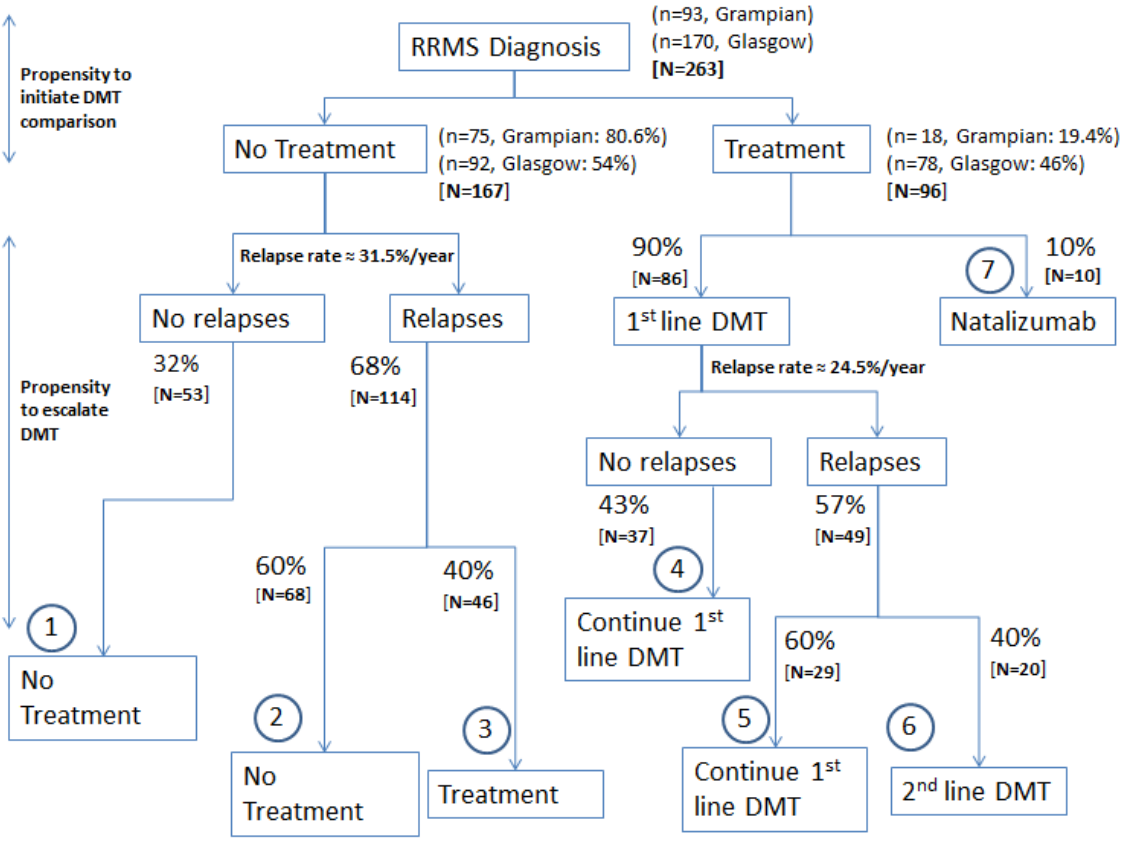
We used data from the literature to estimate the likely disease course of each group over the first 3 years following diagnosis. From pivotal DMT trials, the mean proportion of patients having at least one clinical relapse in the first 2 years ranges from 41-84% (mean 63%) in the placebo arms, to 29-69% (mean 49%) in the treatment arms for low efficacy DMTs and 22-33% (mean 28%) for higher efficacy DMTs. Averaging these figures, we estimated that approximately 63% of the untreated group would have at least 1 relapse in the first 2 years and 28-49% of treated patients would have at least 1 relapse in the first 2 years, depending on which DMT was initiated. The vast majority of patients in the treated group were likely to have been initiated onto platform therapies, and trial populations tend to have milder disease and lower relapse rates, so 49% having at least one relapse in the first 2 years was likely to be the more accurate figure, hence this was used as the estimate and we assumed relapse rates remain relatively constant over time[37]. Trial data over 2 years suggests that 31.5% of patients on no treatment will suffer a relapse each year, and 24.5% of patients on 1st-line DMT will suffer a relapse each year. Cumulatively therefore, we estimated that over a 3-year period approximately 68% of the untreated group (N=167) will have had a relapse (n=114) [See **Table 2-1** and **Figure 2-1**].

For the estimated treated group (N=96), we projected that approximately 90% (n=86) will have been initiated onto a 1st-line DMT than high efficacy treatment (natalizumab)[191]. Approximately 57% (n=49) of the treated group were expected to have a relapse over the first 3 years, assuming a constant relapse rate of 24.5%, and utilising the same calculation process above [See **Table 2-1** and **Figure 2-1**].

Table 2-1: **Projected relapse rates by treatment category (Years 1-3)**

Patients on No Treatment **[N=167 (Annual proportion having a relapse = 31.5%)]**					
	Baseline	**Year 1**	**Year 2**	**Year 3**	**Total by year 3**
Number having relapse	0	**53**	**36**	**25**	**114**
Number relapse-free	167	**114**	**78**	**53**	**53**
Patients on 1st line DMT **[N=86 (Annual proportion having a relapse = 24.5%)]**					
Number having relapse	0	**21**	**16**	**12**	**49**
Number relapse-free	86	**65**	**49**	**37**	**37**

Figure 2-1: *Projected* 3-year treatment pathway and relapse activity for 2010/11 RRMS Cohort



The following assumptions were made to generate these predicted figures:

- Patients who do not suffer relapse activity are unlikely to have any change in their initial management
- Following relapse activity, we estimated 60% of patients on no treatment will remain off treatment and 60% on 1st line DMTs will remain on 1st line treatments
- Patients having tolerability issues with a 1st line DMT are likely to switch to another 1st line treatment rather than escalate, and so will remain in this category

To evaluate the impact of early DMT initiation, we compared outcome for Groups 1 and 2 [(No DMT use (n≈110)] versus Groups 4, 5 and 7 [Early and maintained DMT (n≈70)]. See **Figure 2-1** and **Table 2-2**.

Table 2-2: *Projected* Treatment groups after 3 years

Group		Estimated Total
1	No DMT & no relapse	53
2	No DMT despite relapse	68
3	No initial DMT but escalation to DMT after relapse	46
4	1st line DMT initiated & no relapse	37
5	1st line DMT initiated & continued despite relapse	29
6	1st line DMT initiated & escalation to 2nd line DMT after relapse	20
7	2nd line DMT initiated	10
		263

Using these estimates, a maximum of 70 matched-pairs were therefore possible for analysis. Based on other studies using the propensity matching method[192–195], the propensity-matching process reduces the total cohort number to varying degrees, ranging from 73-93% in the studies quoted, because some patients do not have an adequate match. Thus, the final sample number for outcome comparisons was estimated to be approximately 85% of the cohort, reducing the 70 pairs to approximately 60 matched-pairs for assessment (n≈120).

Data collection

Available clinical records were used to evaluate relevant demographic and clinicoradiological baseline data for each patient, to allow stratification into prognostic groups. Demographic data was available through the SMSR, but detailed information was obtained through hospital patient records. The Community Hospital Index (CHI) number (a unique identifier used throughout Scotland) is stored in the SMSR and provided access to all available clinical data. All records are stored electronically. Data in Greater Glasgow and Clyde were accessed via the Clinical Portal application, whilst records from Grampian were accessed via SciStore©. Formal agreement between NHS GGC and NHS Grampian was provided in order that Grampian records could be accessed remotely for this study.

Each patient's disease course and management were recorded - relapses, disability, DMT use and safety issues, from diagnosis until the present time. Similarly, radiological results are included within the electronic records and used to document MRI findings over time. This established the clinical and radiological baseline status and retrospective disease course profile for all patients.

The OPTIMISE portal was used to capture long-term follow-up data for PS-matched patients. OPTIMISE is a database tool under development in Imperial College London that aims to provide a secure IT framework to capture patient-centred data in MS research and clinical practice. Its aim is to facilitate the capture of prospective longitudinal, standardised clinical and patient-centred data, ideally with an ease that would allow it to be used as an adjunct to routine clinical practice. It will ultimately provide open-source software for management of MS data, integrating anonymised information from multiple centres. This study used the OPTIMISE portal for data capture to pilot the platform before its implementation on a wider scale. The OPTIMISE portal was hosted as an online stand-alone application by NHS Scotland ('Scotland's Health On the Web'), allowing online data entry. The utilisation of OPTIMISE in this way was subject to approval from the Public Benefit and Privacy Panel for Scotland, as outlined in the Ethical Requirements section.

Statistical Analysis

Statistical analysis was undertaken in conjunction with Mario Hair, independent statistics consultant, who calculated the Propensity Scores, generated the matched pairs, produced the population pyramids and undertook outcome analysis for 3-year data using Microsoft Excel® and SPSS®. This section includes some methods which he has written. Longer-term outcome data were analysed using XLStat®.

Hypotheses

Treatment Comparisons

Patients with RRMS in Scotland are managed differently despite similar disease severity (indicated by propensity scoring for DMT initiation).

The null hypothesis is that there are significant differences between mean propensity scores for DMT initiation in those who were treated /escalated and those who were not.

Tests of statistical significance will be undertaken to compare the mean propensity scores, with a p-value of <0.05 considered significant.

Outcome Comparisons

The proportion of patients reaching EDSS 3, suffering relapses and annualised relapse rate will be lower in patients with RRMS initiated onto DMTs, as opposed to those not initiated onto DMTs, but with increased (serious) adverse events.

The null hypothesis is that there is no difference in outcome between patients initiated onto DMTs or not.

Propensity Score Matching

For DMT initiation, we proposed the following factors to estimate the likelihood of starting initial treatment, based on known prognostic factors: these were used in calculating individual Propensity Scores:

- Age
- Gender
- Ethnicity
    - Non-Caucasian
    - Caucasian
- Duration of disease (years since first symptom)
- Type of initial relapse (the presence of each considered ‘high risk’)
    - Motor
    - Sphincter
    - Motor and sensory
    - Polysymptomatology/Multiple neurological systems involved
- Recovery from first relapse
    - Incomplete
    - Complete

- Time from 1st to second relapse (years)
- Number of clinical relapses prior to diagnosis
- MRI at diagnosis
    - Number of T2 lesions
    - Number of Gd-enhancing lesions

Ultimately, the missing data for recovery from first relapse and number of Gd-enhancing lesions at diagnosis was too large to be included in the propensity score matching procedure. In addition, both duration of disease and time from 1st to 2nd relapse were heavily positively skewed so natural log transformations were used on each variable.

The logistic regression model used to estimate the likelihood of starting treatment within the first year (propensity score) therefore used the following covariates:

- Age (in decades) at first diagnosis
- Gender
- Ethnicity
- Natural log of duration of disease
- Initial relapse symptoms
- Natural log of time (years) from 1st to 2nd relapse
- Total number of relapses prior to diagnosis
- Number of MRI brain T2 lesions at diagnosis

Baseline categories for the categorical variables were female, white, zero T2 lesions and sensory only initial relapse.

These were the covariates for the calculation of each patient's Propensity Score (PS) to stratify their likelihood to be prescribed DMTs: those with PSs in the higher strata were also analysed for the frequency of second-line DMTs used initially.

Treatment Comparisons

For comparisons of initial treatment, all available patients with propensity scores for initiation were included in order to provide matches for patients who did or did not start on a DMT within 1 year of diagnosis. The 225 cases with propensity scores were matched using the MatchIt package using a calliper of 0.2. This provided 62 cases from each category matched on propensity score [N=124] (60% of those with no treatment & 51% of those with DMT). The Cohen's d value for standardised difference was used in addition to p-value calculations as a measure of the difference between stated means for some variables. A Cohen's d value of 1 would indicate the two groups differ by 1 standard deviation, a d of 2 indicates 2 standard deviations, etc. As a general guide the difference in means is considered small for Cohen's d values ≤0.2, medium between 0.2 and 0.8 and large if greater than 0.8.

Outcome Comparisons

For outcome comparisons, the trajectory of patients over time had to be known, to ensure only patients who were treated or untreated for the duration were compared, in a bid to replicate treatment and placebo arms in a randomised trial comparing early DMT initiation versus no treatment. Overall, there were 79 patients who had no treatment at all in the first 3 years and 89 patients who had treatment initiated but had no escalation, whether relapses occurred or not. Of these 70/79 and 83/89 had sufficient data to generate a propensity score respectively. The 153 cases with propensity scores were matched using the MatchIt package using a calliper of 0.2. This provided 39 cases from each category matched on propensity score [N=78] (56% of those with no treatment & 47% of those with DMT).

Power calculations

Given the original plan for prospective clinicoradiological review of matched patients, power calculations for demonstrable effect size in the available cohort were made to ensure utility of participant involvement. The estimated sample

sizes are based on the above calculations, reduced by 20% to reflect our estimate that 80% of those patients identified would agree to participate.

Estimated sample size of 96 (48 matched pairs)

Interval level data: A sample size of 48 matched pairs would be sufficient to find a moderate effect size of 0.41 with a power of 0.80. For data such as the mean number of MRI lesions after 6 years an effect size of 0.41 corresponds to a difference of 8 in mean lesions between the treated and untreated groups (assuming the standard deviation of lesions at 6 years is approximately 20).

Categorical data: A sample size of 48 matched pairs would be sufficient to find an odds ratio of 4 with a power of 0.80.

For data such as the proportion reaching EDSS 3.0 over 5/6 years this translates to 20% of the treatment group reaching EDSS 3.0 compared to 50% of the untreated group (assuming half of the matched pairs have different endpoints).

Inclusion Criteria

Adult patients (>16 years)

Diagnosed with RRMS in the Greater Glasgow or Grampian regions between 2010 and 2011

Registered on the SMSR

Exclusion Criteria

PPMS

SPMS

Treatment category definitions

Low efficacy DMTs

- All IFNs
- Glatiramer Acetate (GA)

- Teriflunomide

Moderate efficacy DMTs

- Fingolimod
- Dimethyl Fumarate (DMF)

High efficacy DMTs / treatments

- Natalizumab
- Alemtuzumab
- HSCT (haematopoetic stem cell transplantation)

Efficacy definitions

Relapse

A relapse was defined as the onset of new symptoms or the worsening of pre-existing symptoms attributable to demyelinating disease lasting for more than 24 hours and preceded by improving or stable neurological status for at least 30 days from the onset of the previous relapse, in the absence of infection, fever or significant metabolic disturbance. Objective change on neurological examination was not necessary to fulfill relapse criteria – the occurrence of relapse was at the investigators' (PG) judgment, but where the treating clinician at the time felt a relapse had occurred, this was documented whether meeting the above definition or not.

Relapse severity was judged by a severity score, dividing relapses into the following categories during data collection:

a) ***Moderate*** - a relapse lasting more than 48 hours, fulfilling one of the criteria below:
    - Any motor relapse
    - Any brainstem relapse
    - A sensory relapse if it leads to functional impairment
    - Optic neuritis
    - Intrusive pain

b) ***Severe*** - A disabling relapse is defined as any relapse which fulfils one or more of the following criteria:

- Affects the patient's ability to work
- Affects the patient's activities of daily living as assessed by an appropriate method
- Affects motor or sensory function sufficiently to impair the capacity or reserve to care for themselves or others as assessed by an appropriate method
- Needs treatment (inc. steroids)/hospital admission

Relapses which did not fulfill either of the above criteria were classified as '***Mild***'. A severity score for each relapse was calculated as outlined below also.

Disability

EDSS, where not documented, was estimated from clinical notes using *Neurostatus*® Version 4.0 as a guide. The OPTIMISE database generates an estimated EDSS score based on examination findings at each clinic visit but this was subject to approval during data entry.

Ethical Considerations

MODERATE was approved by the West of Scotland Research Ethics Committee (Ref 16/WS/0017). Two further substantial amendments were also approved, the second being approval of the retrospective phase of the original study as a stand-alone project, as presented here, following the funding issues outlined above. The study was ultimately funded via a local NHS Endowment Fund.

The study was a retrospective, observational assessment of clinical records in two regional MS centres in Scotland and did not require subject participation. Access to clinical records for the purposes of this study was entirely limited to the study team. Cross-board patient data access was required and was agreed across sites and in line with national regulations, as per centre-specific IT/data access guidelines. Data was entered on NHS-networked devices on NHS premises in Glasgow. 'Scotland's Health On the Web' (www.show.scot.nhs.uk) hosted the OPTIMISE database on a secure NHS server in order to permit remote data entry

and storage from each site. This required password-protected access in order to enter data. There was no transfer of data outwith the two regions involved, but storage was centralised. Data were stored anonymously, using unique identifiers, on NHS-approved storage devices and backed up on local servers.

An Information Request Form was submitted to Scottish Healthcare Audits, part of the Information Services Division of NHS Scotland, to access the SMSR data. The decision to release the data was made by the Research Subgroup of the SMSR Steering Group, of which Dr Overell is a member. Approval to access patient data, using the CHI number stored on the SMSR, was sought from both Caldicott Guardians in Glasgow and Grampian and the study was approved by the Public Benefit and Privacy Panel for Scotland, with regards to information governance.

Results

Cohort Characteristics

A total of 397 patients were registered on the SMSR with a new diagnosis of MS from the Greater Glasgow and Grampian regions between 2010 and 2011. Clinical records were available for 375 patients, 301 (80.3%) of which were diagnosed with RRMS, 57 (15.2%) PPMS and 17 (4.5%) SPMS (See **Figure 2-2**, **Tables 2-3** and **2-4**).

Figure 2-2: MODERATE cohort disposition from SMSR

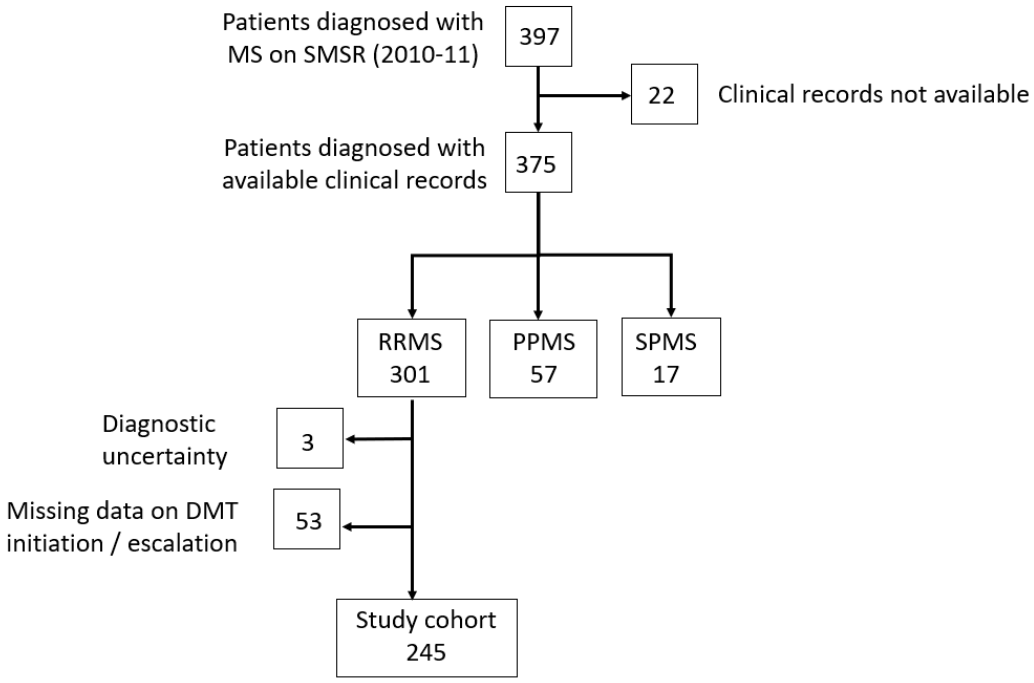


Table 2-3: Number of patients registered and MS subtype on SMSR 2010-2011

Diagnosis	N (%)	Year		Location	
		2010	2011	Grampian	GGC
RRMS	**301 (80.3)**	163	138	86	215
PPMS	**57 (15.2)**	32	25	15	42
SPMS	**17 (4.5)**	14	3	8	9
Total	**375**	222	175	119	278

Table 2-4: Number of patients on SMSR 2010-2011 - details

Region	No. of patients														
	Total*			RRMS			PPMS			SPMS			Other		
	2010[a]	2011[b]	**Total[c]**	2010	2011	**Total**	2010	2011	**Total**	2010	2011	**Total**	2010	2011	**Total**
GGC	102	88	**190**	76	73	**149**	19	10	**29**	6	2	**8**	1	3	**4**
Lanarkshire	43	26	**69**	37	18	**55**	5	7	**12**	1	0	**1**	0	1 (No info)	**1**
Ayrshire	5	1	**6**	4	1	**5**	1	0	**1**	0	0	**0**	0	0	**0**
Western Isles	4	2	**6**	4	2	**6**	0	0	**0**	0	0	**0**	0	0	**0**
Grampian	63	56	**119**	42	44	**86**	7	8	**15**	7	1	**8**	7	3	**10**
Total	217	173	**390**	163	138	**301**	31	25	**57**	14	3	**17**	8	7	**15**

*Patients excluded as no information available (Non-Grampian patients only)

a = 5 (4 invalid CHI provided, 1 no information on Portal)

b = 2 (Invalid CHI)

c = 7

Of the 301 patients with RRMS, 245 patients had appropriate data available within their clinical records on DMT initiation and escalation to be included in the final cohort. Of the 53 patients excluded due to missing data, 28 were from Grampian [19% of the total (N=147) Grampian cohort] and 25 [8.3% of the total (N=303) Glasgow cohort] hence proportionally more Grampian patients were excluded. Of the 3 patients excluded due to diagnostic uncertainty, 2 were from Glasgow and 1 from Grampian.

The cohort was fairly typical for a population of patients with RRMS (See **Table 2-5**). Predominantly female, the cohort was also almost entirely Caucasian, and mean 36 years of age at diagnosis. The median duration since first symptom at diagnosis was 1.8 years but there was a wide range from weeks to decades; similarly, this was also the case for the duration between first and second relapses, but on average this was just over a year. Notably, given the retrospective nature of the analysis, relapses which occurred after diagnosis were included in calculating the duration between first and second relapses as not all patients had had 2 relapses at diagnosis. The vast majority (223/245, 91%) had at least 2 relapses prior to diagnosis, however. Data was incomplete for recovery from first relapse but, of those with data available, the majority recovered completely although over a fifth of patients had incomplete recovery even after 6 months. Baseline MRI brain data were available for the majority of patients and most (64.5%) had at least 9 T2 lesions.

Of the 245 patients included, 130 (53%) had a DMT initiated within one year of diagnosis while 115 (47%) did not. The proportion of patients treated with a DMT at diagnosis was notably lower than we had predicted prior to the study, as we anticipated over 60% of the cohort would be treated within 1 year (See **Figure 2-1**). Indeed, it is immediately surprising that the chance of starting on a DMT in this cohort was almost 50:50 at the outset, implying virtually random chance of being treated or not irrespective of disease status. However, 53% of patients receiving a DMT in our cohort is higher than the average overall treatment rates in Scotland purported in the survey upon which we based our projections (36%)[190] and the international MS atlas estimates for the UK as a whole (11%)[196]. More recent analysis suggests that around 59% of eligible patients in the UK have access to DMTs however[197]. Importantly, we did not analyse differences between the two regional centres directly.

The vast majority of patients started on DMTs within 1 year of diagnosis used injectable therapies, as shown in **Figure 2-3**. Overall, 118/130 (91%) used either IFN or GA first-line. One patient was started on Azathioprine as a dual treatment for MS and retinal vasculitis. Just 11 patients (8.5%) started directly on 2nd-line therapy, the majority of which used natalizumab (9/11), with alemtuzumab and mitoxantrone used by the others.

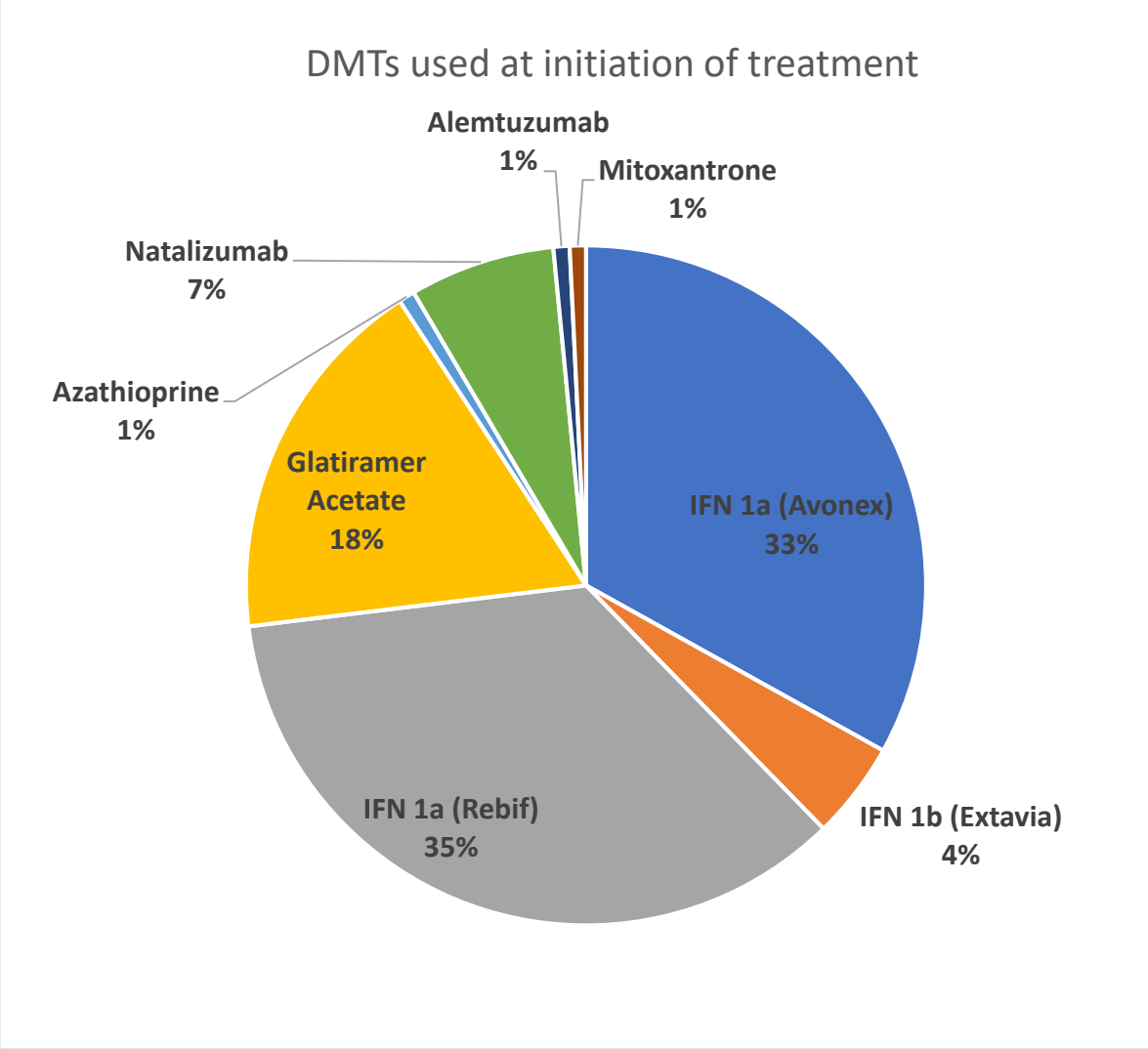
More detailed comparisons of those initiated onto a DMT after diagnosis or not will be presented but there are notable differences at the outset. Treated patients were more likely to be female and were, on average, significantly younger than those who did not receive treatment (See **Table 2-5**). Duration of disease at diagnosis and the time between first and second relapses was also notably shorter in the treated group. Additionally, a higher proportion had had 4 or more relapses prior to diagnosis in those who went on to have treatment but, notably, this was also the case for 4 patients (3.5%) within the untreated group. Indeed, almost a quarter of the untreated group had 3 or more relapses prior to diagnosis, albeit these were more likely to be spread over a longer time. Recovery from first relapse and MRI lesion load does not appear to have been a major factor in treatment decisions, but a higher proportion of the treated group had ≥9 MRI brain T2 lesions in comparison to those who did not start treatment. Together, these data suggest that young female patients with short disease duration and frequent relapses close together were prioritised for DMT initiation in our cohort.

Table 2-5: MODERATE Cohort characteristics

	Total cohort	DMT started <1yr	DMT not started <1yr
N (%)	245 (100)	130 (53)	115 (47)
Female (%)	168 (69)	92 (71)	76 (66)
Mean age at diagnosis (SD) [years]	36 (9.8)	33 (9.2)*	39 (9.6)
Median duration of disease at diagnosis (Range) [years]	1.80 (0.05-29.7)	1.32 (0.05-21.3)*	3.10 (0.15-29.7)
Median duration between 1st and 2nd relapses (range) [years]	1.19 (0.04-20.29) [n=238]	0.66 (0.04-20.29)* [n=128]	2.87 (0.08-19.55) [n=110]
Relapses before diagnosis			
Mean (SD)	2.28 (0.75)	2.35 (0.83)	2.20 (0.64)
Number	n (%)		
1	21 (8.6)	11 (8.5)	10 (8.7)
2	154 (62.9)	78 (60)	76 (66.1)
3	51(20.8)	26 (20)	25 (21.7)
4	16 (6.5)	12 (9.2)	4 (3.5)
5	2 (0.8)	2 (1.5)	0
N/D	1 (0.4)	1 (0.8)	0
Recovery from 1st relapse	n (%)		
Complete at <3/12	93 (38.0)	53 (40.8)	40 (34.8)
Incomplete at 3/12	33 (13.5)	21 (16.2)	12 (10.4)
Incomplete at 6/12	50 (20.4)	27 (20.8)	23 (20)
N/D	69 (28.2)	29 (22.3)	40 (34.8)
Number of MRI Brain T2 lesions at diagnosis			
0	8 (3.3)	4 (3.1)	4 (3.5)
<9	69 (28.2)	34 (26.2)	35 (30.4)
≥9	158 (64.5)	89 (68.5)	69 (60)
N/D	10 (4.1)	3 (2.3)	7 (6.1)

*p<0.05

Figure 2-3: DMTs used at initiation of treatment (n=130)



Treatment comparisons

The following primary endpoints will be addressed in this section

- Mean DMT initiation propensity scores in patients initiated onto a DMT within 1 year of diagnosis and those not initiated onto a DMT within 1 year of diagnosis

There was a total of 20 missing cases so the model estimated propensity scores for 225 patients (103/115, 90% of those with no DMT & 122/130, 94% of those with DMT). The Cox & Snell r^2 for the model was a modest 0.212 and only age, time between relapses and number of relapses had p values of less than 0.2; this suggests the model was relatively weak at predicting the outcome variable but

these factors were most important. However, analysis comparing actual outcomes with the propensity scores i.e. likelihood of being in the chosen treatment group based on the propensity score calculated.

The mean propensity score (SD) for treatment was significantly higher in patients initiated onto a DMT within 1 year of diagnosis [0.64 (±0.19)], in comparison to those who did not start on a DMT [0.42 (±0.23), $p<0.001$]. This reflects patient and clinician decisions to start treatment in those with more severe disease. However, **Figure 2-4** shows the propensity scores for treatment distributed across the entire cohort and, whilst those who started treatment within a year generally had higher scores than those who did not, there is considerable overlap. **Table 2-6** shows DMT initiation by propensity score quartile, with 42% (43/103) of those having no treatment and 57% (70/122) of those starting treatment in the middle two quartiles and hence likely comparable baseline disease. These data suggest that, whilst on average patients with worse prognostic factors at diagnosis are more likely to receive treatment, a significant proportion of patients are being treated differently despite comparable disease severity.

It is also notable that patients with propensity scores at the extremes of the cohort were not necessarily managed as might be expected. For example, 12/56 (21%) of the patients in the highest PS quartile for DMT initiation did not start on a DMT within 1 year and, of those in the lowest PS quartile for DMT initiation (n=56), 9 (16%) were commenced on a DMT within 1 year of diagnosis. This suggests that more patients are being under- rather than over-treated, perhaps. Interestingly, of those initiated directly onto a 2nd line DMT (natalizumab) [Group 7, n=11], 6 (55%) were in the highest quartile (4), 3 were in the 3rd quartile and 1 was in the second quartile (1 did not have enough available data to calculate a PS) and none from the first quartile, suggesting this group of patients were appropriately initiated onto high efficacy treatment from the outset.

Figure 2-4: Population pyramid showing distribution of propensity scores by group

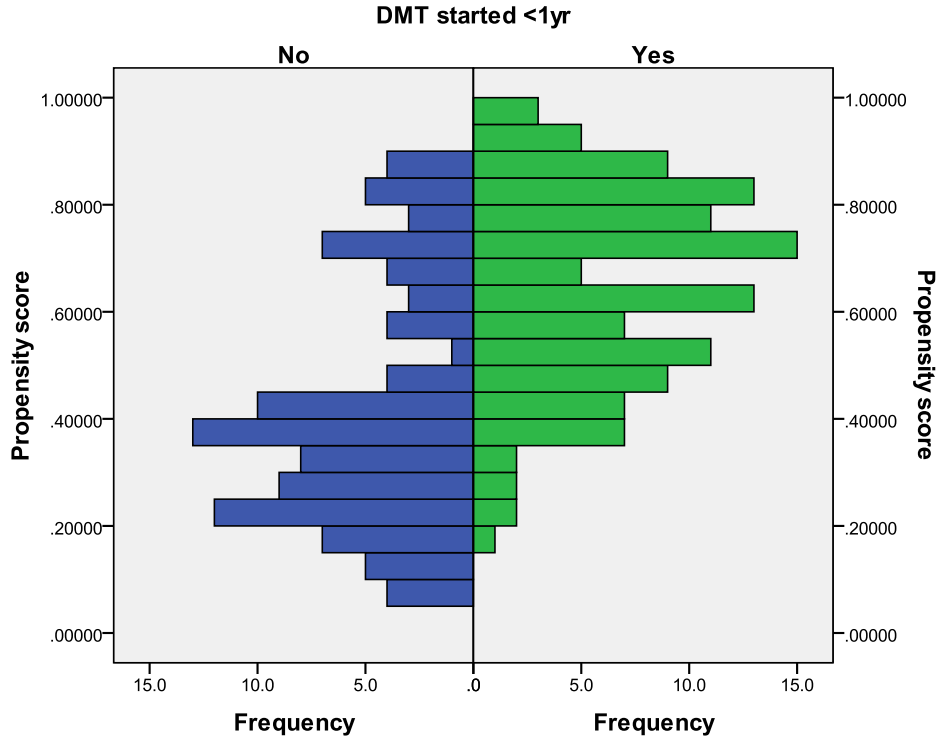


Table 2-6: Proportion of DMT category by propensity score quartile

Propensity score quartiles	DMT started <1yr?		Total
	No	Yes	
1 (lowest likelihood of receiving treatment)	47 (46%)	9 (7%)	56
2	26 (25%)	30 (25%)	56
3	17 (16%)	40 (33%)	57
4 (highest likelihood of receiving treatment)	13 (13%)	43 (35%)	56
Total	103	122	225

Using PSM, we were able to match 124 (55%) of the 225 patients on prognostic factors at diagnosis and there was no longer a significant difference in the mean propensity score for treatment despite half the group being treated and half not (62 matched pairs). **Table 2-7** shows the cohort characteristics before and after the matching process. In the unmatched cohorts, patients receiving a DMT within 1 year were significantly younger at diagnosis, with a shorter duration of disease

and inter-relapse interval. The cohorts were otherwise comparable but those starting a DMT within 1 year had a significantly higher mean PS for treatment. After matching, there was no significant difference in disease characteristics, as reflected in the comparable mean PS for treatment, yet they were treated differently. This suggests factors other than the disease characteristics included in our propensity score are influencing treatment decisions. These may be patient- or clinician-related but, notably, only 8 (7%) of the 115 patients who didn't start a DMT within 1 year of diagnosis were documented to have refused treatment for imminent pregnancy plans (6, 5%) or other reasons (2, 2%). As outlined below, however, follow-up information was limited for many patients and incomplete documentation may result in this being an underestimate, but it does not appear to be the case that the vast majority of patients refused treatment where it was offered or suggested.

The propensity score using prognostic factors at baseline was notably predictive of treatment within the first 3 years, being lowest in those receiving no DMTs and having no subsequent relapses and highest in those initiated directly onto 2nd-line therapy, as shown in **Table 2-8** and **Figure 2-5**. Using a Tukey B post-hoc test, mean PS was significantly higher in those starting and escalating DMTs in comparison to those who remained untreated in the first three years. This validates the predictive value of our model although the confidence intervals overlap outwith the extremes.

Table 2-7: Baseline characteristics of patients before and after matching

	Unmatched			Matched		
	DMT started <1yr			DMT started <1yr		
	No	Yes	d†	No	Yes	d†
N (% female)	103 (65%)	122 (70%)		62 (71%)	62 (66%)	
Age at diagnosis (mean±SD)	38.3±9.6	33.5±9.2	**0.51***	35.1±9.2	35.8±9.1	0.08
Duration of disease (years)	5.7±5.7	2.7±4.1	**0.60***	3.5±4.5	3.7±4.7	0.06
Years between relapse	4.3±4.3	1.6±2.7	**0.76***	2.4±3.3	2.4±3.5	0.001
Number of relapses prior to diagnosis	2.2±0.7	2.4±0.8	0.21	2.2±0.7	2.5±0.9	0.32
White ethnicity (%)	102 (99%)	118 (97%)		61 (99%)	62(100%)	
Initial relapse symptoms						
Sensory only	68 (66%)	77 (63%)		42 (68%)	43 (69%)	
Motor	15 (14%)	16 (13%)		6 (10%)	6 (10%)	
Sensorimotor	12 (12%)	15 (12%)		8 (13%)	5 (8%)	
Polysymptoms	7 (7%)	12 (10%)		5 (8%)	8 (13%)	
Sphincter involvement	1 (1%)	2 (2%)		1 (1%)	0 (0%)	
Number of T2 lesions on Baseline MRI Brain						
0	4 (4%)	3 (3%)		3 (5%)	1 (2%)	
<9	33 (32%)	33 (27%)		21 (34%)	21 (34%)	
≥9	66 (64%)	86 (70%)		38 (61%)	40 (64%)	
Propensity score	0.42±0.23	0.64±0.19	**1.07***	0.54±0.21	0.57±0.20	0.13 p=0.46
Recovery from 1st relapse	n=72	n=97		n=46	n=49	
Complete	39 (54%)	51 (53%)		24 (52%)	21 (43%)	
Partial (incomplete at 3/12)	11 (15%)	20 (20%)		8 (17%)	10 (20%)	
None (incomplete at 6/12)	22 (31%)	26 (27%)		14 (30%)	18 (37%)	

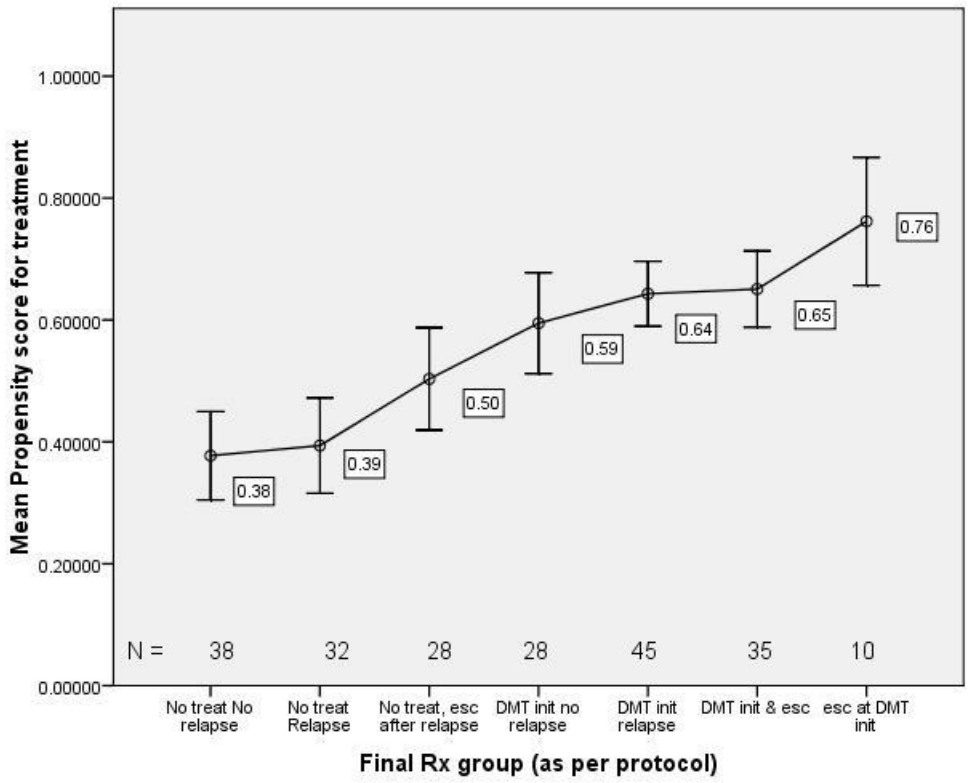
†Standardised difference (Cohen's d) for continuous variables.

*p < 0.001

Table 2-8: Patient groupings after 3 years from diagnosis

Group	Outcome at 3 years	DMT started <1yr		Mean Propensity Score		Total
		No	Yes	DMT initiation (n)	DMT escalation (n)	
1	No DMT & no relapse	42 (37%)	0	0.38 (38)	0.21 (41)	42
2	No DMT despite relapse	37 (32%)	0	0.39 (32)	0.27 (34)	37
3	No initial DMT but escalation to DMT after relapse	31 (27%)	0	0.50 (28)	0.37 (28)	31
4	1st line DMT initiated & no relapse	0	30 (23%)	0.59 (28)	0.26 (27)	30
5	1st line DMT initiated & continued despite relapse	0	48 (37%)	0.64 (45)	0.33 (46)	48
6	1st line DMT initiated & escalation to 2nd line DMT after relapse	0	37 (29%)	0.65 (35)	0.54 (34)	37
7	2nd line DMT initiated	0	11 (9%)	0.76 (10)	0.48 (10)	11
8	Escalation after MRI not relapse	5 (4%)	4 (3%)	0.56 (9)	0.23 (9)	9
	Total	**115**	**130**	**0.54 (216)**	**0.33 (229)**	**245**

Figure 2-5: Mean propensity score (95% CIs) at diagnosis and treatment group after 3 years



Matched cohort outcome comparisons

The following secondary endpoints will be addressed in this section:

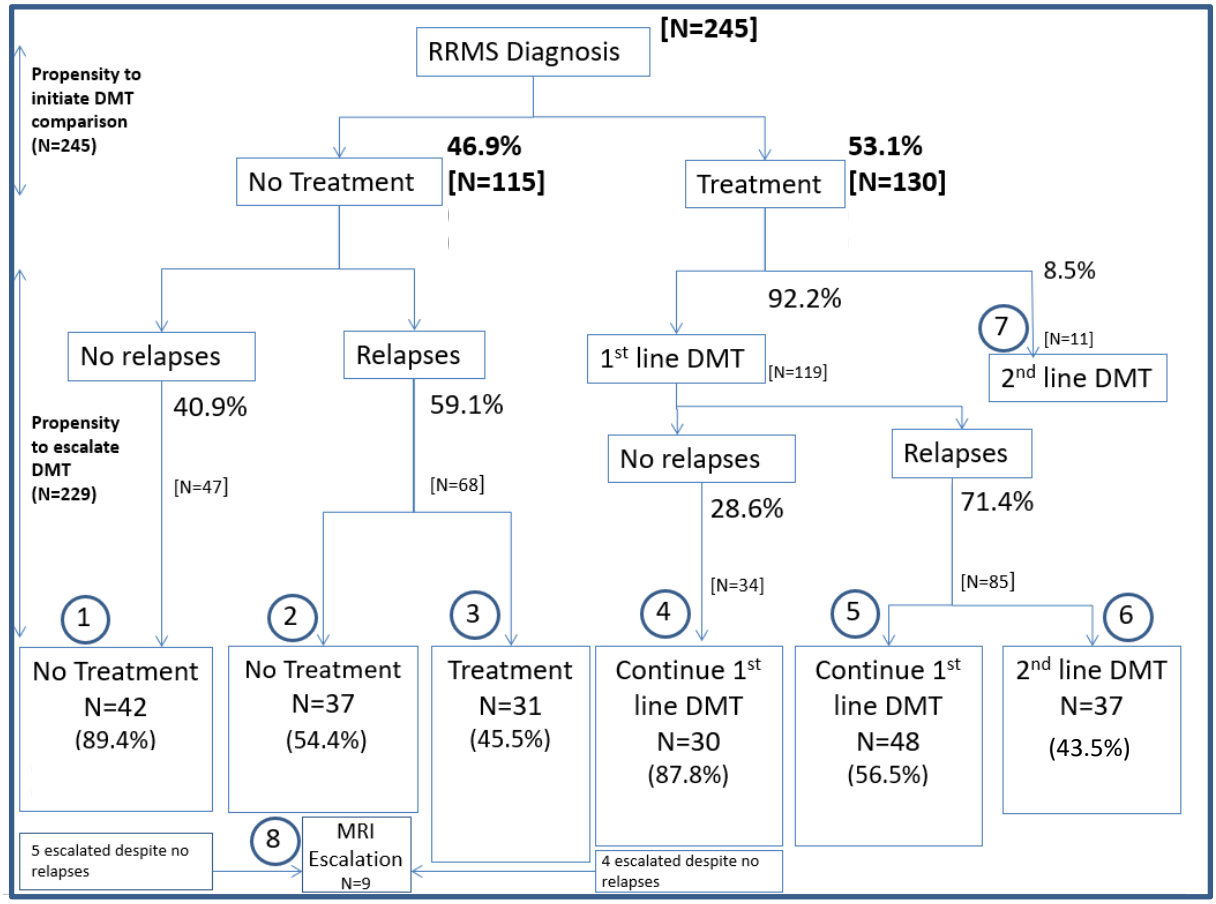
- Proportion of patients reaching EDSS 3 in first 5-6 years after diagnosis in each group
- Proportion of patients relapsed in first 5-6 years after diagnosis in each group
- Annualised relapse rate in first 5-6 years after diagnosis in each group
- Proportion experiencing (serious) adverse events in first 5-6 years after diagnosis in each group

In order to make meaningful outcome comparisons, we evaluated treated and untreated PS-matched cohorts over time. This included analysis of both 3- and 6-year outcomes where available. However, there was significantly less available follow-up data available for the untreated cohorts who were often discharged from clinic, meaning longer-term outcomes are skewed towards identifying events in treated patients, simply because they were being evaluated whilst

untreated patients were not. This is a significant bias that has to be borne in mind when interpreting comparative outcomes.

The following sections describe early (3-year) and longer-term (6-year) clinical outcomes comparing treated and untreated patients. The patients were grouped by treatment and clinical outcome at the end of available follow-up as outlined in **Table 2-8** and **Figure 2-6** for the entire cohort.

Figure 2-6: Patient disposition in first 3 years after diagnosis



- **Group 1**: No DMT initiation and no relapse activity during follow-up
- **Group 2**: No DMT initiation despite relapse activity during follow-up
- **Group 3**: No DMT initiation in <1 year from diagnosis but escalated to treatment after relapse activity during follow-up

- **Group 4**: 1st-line DMT initiated in <1 year from diagnosis and no relapse activity during follow-up

- **Group 5**: 1st-line DMT initiated in <1 year from diagnosis and continued despite relapse activity during follow-up

- **Group 6**: 1st-line DMT initiated and escalated to 2nd-line DMT after relapse during follow-up

- **Group 7**: 2nd-line DMT initiated in <1 year from diagnosis

- **Group 8**: Escalation from no DMT or 1st-line DMT to 1st- or 2nd-line DMT respectively after MRI activity alone

There were 39 matched pairs (N=78) of patients either treated with a DMT <1 year after diagnosis or not. The untreated patients included Groups 1 & 2 (no DMT ever whether relapsed or not) and the treated patients comprised Groups 4, 5 & 7 (started DMT <1 year after diagnosis and continued whether relapses or not but not escalated). **Table 2-9** shows the baseline characteristics for the cohorts before and after matching. Similar to the overall cohort, the treated group were significantly younger, with shorter duration of disease and inter-relapse interval in comparison to untreated patients in the unmatched cohorts. There were no other statistically significant differences. The higher levels of disease activity are reflected in the significantly higher PS for initiation in treated patients. However, there is no significant differences in the baseline characteristics after matching, including PS for initiation, implying that, while this subgroup reflects the overall cohort where patients with similar disease activity are managed differently, these cohorts had comparable prognosis at diagnosis hence differences in outcome are considered related to treatment effects.

Table 2-9: Baseline characteristics of DMT initiation patients for 3-year outcome comparisons (before and after matching)

	Unmatched			Matched		
	DMT started <1yr			DMT started <1yr		
	No	Yes	d†	No	Yes	d†
N (% female)	70 (64%)	83 (69%)		39 (72%)	39 (62%)	
Age at diagnosis (mean±SD)	39.6±9.1	34.9±8.4	**0.54***	36.5±9.5	36.9±8.7	0.04
Duration of disease	6.0±6.0	2.3±3.4	**0.77***	3.5±4.1	3.4±4.4	0.01
Years between relapse	4.7±4.6	1.2±1.5	**1.05***	2.2±2.5	1.8±1.9	0.19
No of relapses	2.1±0.6	2.3±0.8	0.17	2.2±0.6	2.1±0.7	0.08
Ethnicity (% white)	70 (100%)	81 (98%)		39(100%)	39(100%)	
Initial relapse symptoms						
Sensory only	47 (67%)	56 (68%)		28 (72%)	25 (64%)	
Motor	10 (14%)	10 (12%)		5 (13%)	6 (15%)	
Sensorimotor	9 (13%)	9 (11%)		4 (10%)	4 (10%)	
Polysymptoms	3 (4%)	7 (8%)		1 (3%)	4 (10%)	
Sphincter involvement	1 (1%)	1 (1%)		1 (3%)	0 (0%)	
Number T2 lesions on Baseline MRI Brain						
0	3 (4%)	2 (2%)		1 (3%)	1 (3%)	
< 9	24 (34%)	22 (27%)		16 (41%)	10 (26%)	
≥9	43 (61%)	59 (71%)		22 (56%)	28 (72%)	
Propensity score	0.38±0.22	0.64±0.19	**1.26***	0.51±0.21	0.54±0.20	0.13 p=0.58
Recovery from 1st relapse	n=50	n=66		n=31	n=32	
Complete	27 (54%)	32 (49%)		17 (55%)	15 (47%)	
Partial (incomplete at 3/12)	5 (10%)	15 (23%)		3 (10%)	5 (16%)	
None (incomplete at 6/12)	18 (36%)	19 (29%)		11 (35%)	12 (37%)	

†Standardised difference (Cohen's d) for continuous variables.

*p < 0.001

Clinicoradiological outcome measures were evaluated retrospectively using medical records from 2010-2017. Notably, these patients were matched on their propensity score for DMT initiation but in the knowledge of their outcomes at 3 years from diagnosis, allowing their allocation into Groups 1-7 and, hence, meaningful group comparisons (See **Figure 2-6**). For 6-year outcomes, patient disposition was similarly categorised after 6 years, as shown in **Figure 2-7**.

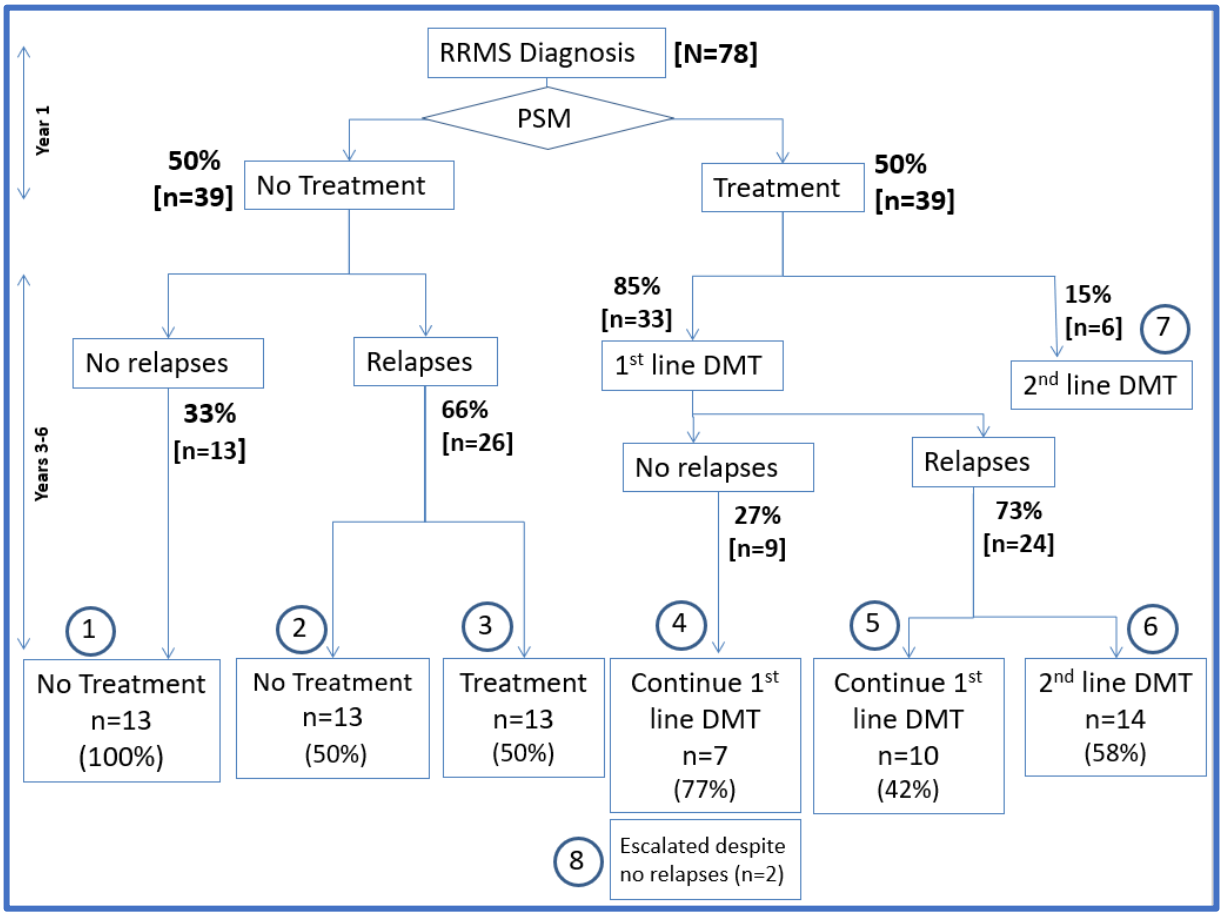
Three analyses of outcomes at 6 years from diagnosis are presented (Groupings refer to **Figure 2-7**):

- Entire matched cohort: DMT started < 1yr or not (irrespective of DMT use after 3 years)
    - N =78 (39 started DMT <1year vs 39 no DMT in first 3 years)
    - Groups 4-8 vs Groups 1-3

- Early DMT initiation versus late DMT initiation
    - N = 36 (23 started DMT <1year vs 13 started DMT after 3 years, after relapse)
    - Groups 4,5 & 7 vs Group 3

- Early DMT initiation (+/- escalation) vs. No DMT
    - N = 65 (39 started DMT <1year vs 26 no DMT)
    - Groups 4-8 vs Groups 1&2

Figure 2-7: Patient disposition in first 6 years after diagnosis [N=78]



Entire Matched Cohort

Firstly, outcomes after approximately 6 years for the entire matched cohort (N=78) are presented in **Table 2-10**. This represents the effects of early (<1 year) DMT initiation where patients were treated or untreated for the first 3 years. A third of the untreated group were however treated with DMTs after 3 years (Group 3), but started treatment a median of over 4 years after diagnosis and the proportion of time spent on a DMT was, on average, less than 1 year over the follow-up period. **Figures 2-8 and 2-10** summarise all DMTs to which the cohorts were exposed over the study period, whilst **Figures 2-3 and 2-9** outline the initial DMTs used in each cohort: notably, the oral therapies were not available at the time the early treatment group started treatments and non-MS immunomodulatory therapies [e.g. Mycophenylate (MMF) and Tacrolimus] were used in some patients with additional non-MS autoimmune diseases as dual therapy. These data demonstrate that injectable therapies were the most frequent 1st-line DMTs used and remained the most widely used throughout

follow-up. There is notable use of oral therapies though (teriflunomide, dimethyl fumarate and fingolimod), with a higher proportion of these used by those starting treatment after 3 years. However, high efficacy treatments (natalizumab, alemtuzumab and mitoxantrone) were *only* used in patients starting treatment early. **Figure 2-11** illustrates the differences in DMTs used by patients initiated onto treatment earlier or later. These data establish that, when DMTs were used, most patients were exposed to injectable therapies whether treatment was started early or late but oral DMTs were used proportionally more in those starting treatment later and high efficacy infusion therapies only in those starting earlier, albeit in small numbers.

The outcomes for the entire matched cohort outlined in **Table 2-10** were intended to reflect the impact of early DMT initiation, but, in practice, the follow-up duration was significantly longer in treated than untreated patients which biases the results towards capturing increased events in the treated cohort. Whilst on average patients were diagnosed over 6 years before the end of data collection, potentially allowing 6-year outcome comparisons as planned a priori, treated patients had available follow-up correspondence for 5.2 years on average and untreated patients just 3.3 years. Additionally, the treated cohort had a higher number of visits, documented EDSS scores and MRI scans undertaken over this longer period, permitting increased capture of events in comparison to the untreated group. It is also notable that baseline EDSS score was significantly higher in the treated group. This was not a covariate in the propensity score for DMT initiation and hasn't been balanced in the matching process, but implies the treated group were more disabled at the outset and this may also impact on outcomes, unrelated to treatment decisions, as higher baseline EDSS is known to be a strong predictor of worse outcome.

With these caveats in mind, efficacy and safety outcomes can be considered. By definition, the treated group had a longer duration of DMT treatment during follow-up in comparison to the initially untreated group (4.2 vs 0.6 years). Additionally, the treated group used over 2 DMTs on average during follow-up and were initiated onto a DMT within approximately 4 months of diagnosis. Outwith DMTs, the earlier treated group also used more symptomatic therapies than those not using DMTs initially (2.5 vs 1.1 treatments per patient). This likely simply reflects the longer follow-up and, hence, access to symptomatic

treatments, but may also represent a higher burden of symptoms. The category of symptomatic treatments used is outlined in **Figure 2-12**. Analgesia, predominantly neuropathic, was the most frequent indication for symptomatic therapy and in all categories except (male) sexual dysfunction, patients starting a DMT earlier had more treatments.

Table 2-10: Entire matched cohort 6-year outcomes for DMT initiation

	DMT started <1yr	No DMT started <1yr	p
Baseline characteristics			
N (% female)	39 (62)	39 (72)	
Mean age at diagnosis (SD)	36.9 (8.7)	36.5 (9.5)	0.99
% Female	62	72	0.5 (z-test)
Mean no. Co-morbidities	1.33	1.46	0.69
Mean disease duration at diagnosis (SD) [years]	3.4 (4.4)	3.5 (4.1)	0.8
Mean Baseline EDSS	2.7* (Median 2.5)	1.7 (Median 1.5)	0.01 (Mann-Whitney 0.026)
Baseline number of T2 lesions (%) <9 >9 0	(n=34) 31.3 65.6 3.1	(n=36) 41.7 55.6 2.8	z-tests 0.4 0.77 1
Follow-up			
Mean duration of follow-up (yrs)	5.2*	3.3	0.0004
Mean Number of visits per patient	7.7*	4.4	0.0001
Mean duration from diagnosis to end of data collection	6.3	6.3	
Number of visits per year of study	1.2*	0.7	$0.7x10^{-5}$
Mean number of visits per year of follow-up	1.7	2.1	0.5
Treatments			
Mean duration on any DMT (years)	4.2*	0.6 (n=13)	$0.48x10^{-14}$
Mean number of DMTs used	2.1*	0.6	$0.95x10^{-8}$
Mean duration from diagnosis to DMT commencement (years)	0.3* (median 0.3)	3.6 (median 4.4)	0.00015
Total number of symptomatic treatments used	96	44	

Mean number of symptomatic treatments used per patient	2.5*	1.1	0.004
Relapses			
Total number of relapses recorded after diagnosis	62	51	
Proportion with 0 relapses after diagnosis	12/39	13/39	1.0
Mean total number of relapses per patient after diagnosis	1.6	1.3	0.4
Mean duration to 1st relapse after diagnosis (years)	1.8	1.5	0.6
Mean ARR during follow-up	0.3	0.7	0.09
Mean ARR (excluding relapses <1yr after diagnosis)	0.3	0.4	0.17
Severity of relapses (%)			
Mild	34.9	52.9	0.08 (z)
Moderate	38.7*	11.8	0.001
Severe	6.6	2.0	0.449
Unknown	19.8	33.3	0.159
Steroids used	35.5	9.8	0.001
EDSS			
Total EDSS scores recorded	158 (n=37)	89 (n=37)	
EDSS scores documented per patient	4.3	2.4	
Mean EDSS scores per patient per year of follow-up	0.9 (Median 0.8)	1.7 (Median 0.6)	0.2
Mean change from Baseline EDSS to final EDSS	+1.1	+0.86	0.75
Mean duration between baseline and final EDSS (years)	5.2 (Median 6.0)	3.9 (Median 4.4)	0.054
Proportion developing SPMS	8/39	2/39	0.08 (z-score)
Proportion reaching EDSS 3 (maintained until	17/33	9/23	

last available EDSS) [denominator is number with ≥2 EDSS scores]			0.5 (z-score)
MRI			
Total number of MRIs undertaken [inc contrast]	141 [79] (n=34)	84 [21] (n=36)	
Number of patients having ≥1 MRIs with contrast	30	11	
Mean number of MRIs per patient	3.6*	2.2	0.004
Mean MRIs per patient per year of follow-up	0.8 (Median 0.6)	1.3 (Median 0.6)	0.385
Patients developing new or enlarging T2 lesions	19/34	13/36	0.15 (z-test)
Patients developing new Gd lesions	11/30 (37%)	5/11 (45%)	0.65 (z-test)
Treatment-related adverse events			
Number of patients having AE leading to Rx discontinuation	21 (53.8%)	8/13 (61.5%)	
DMT-related AEs - total	34	10	
Mean number of AEs leading to Rx discontinuation per patient	0.9*	0.3	0.014
SAEs	Death from PML		

*p<0.05

Relapses

There was no statistically significant difference in relapse frequency but steroid use and severity significantly favoured the initially untreated group. Despite the difference in duration of follow-up, the number of relapses recorded in total was similar for both groups. However, relapses had occurred in around two thirds of each group, which favours the early treated group given the longer follow-up,

assuming more events would have occurred in the untreated group had they been evaluated. Additionally, the number of relapses per patient (1.6 vs 1.3), time to 1st relapse (1.8 vs 1.5 years) and ARR (0.3 vs 0.7) all favoured the early treatment group numerically but there was no statistically significant difference in these outcomes. In an attempt to account for immortal time bias with relapses, only those occurring after at least 1 year after diagnosis in both groups were also evaluated and this resulted in more comparable absolute ARRs but remained no statistically significant difference. A significantly higher proportion of the untreated group had milder relapses and less need for steroids but there was missing data in both cohorts for these measures and proportionally more in the untreated group although this difference was not statistically significant. Overall, these data may suggest favouring early treatment for relapse outcomes in terms of absolute numbers and given the longer follow-up, but the lack of statistical significance does not permit this as a conclusion.

Disability

As outlined above, the earlier treated group had a significantly higher baseline EDSS and longer follow-up but there was no statistically significant difference in disability outcomes between the groups. EDSS scores were documented in 37/39 patients from each cohort, but only 33 from the early treatment group and 23 from the later/no treatment group had more than two scores documented to allow comparisons over time. The number of EDSS scores documented was higher in the early treatment group but the untreated group had a higher frequency of recordings during their shorter follow-up. The early treated group had approximately 1 EDSS score documented per year of follow-up whilst the untreated group had almost 2 per year. There was a higher absolute increase from baseline to final EDSS score in the early treated group (+1.1 vs +0.86) but this was not statistically significant and, again, the final EDSS was recorded with over a year of additional disease duration in the early treated group in comparison to the initially untreated cohort. Proportionally, more patients in the early treated group were documented as developing SPMS (8/39 vs 2/39) and reached EDSS 3 (17/33 vs 9/23) but these differences were not statistically significant. These data suggest the early treated group developed more disability over time but this is not statistically significant and likely influenced by factors

other than DMT use, namely the higher baseline EDSS and longer duration of follow-up.

MRI

As with the outcomes above, the early treatment group had more MRI evaluation over a longer period but no statistically significant differences occurred in terms of lesion development between the two cohorts. In the early treatment group, 34 patients had at least 1 MRI brain and a total of 141 scans were undertaken. Intravenous contrast was used in 79/141 MRIs and 30 patients (77%) had at least one contrast-enhanced scan. In the untreated group, however, only 84 MRI brain scans were undertaken in 36 patients and 21 of the scans included contrast-enhancement, undertaken in just 11 patients (28%). This meant patients in the early treatment group had almost twice as many scans in total (p=0.004), albeit the frequency of scanning was higher in the untreated group during their shorter follow-up, but this difference was not statistically significant. There was no statistically significant difference in lesion load at baseline between the groups but a higher proportion of the early treated group [19/32 (56%)] developed new or enlarging T2 MRI brain lesions in comparison to the untreated group [13/36 (36%)] although this is likely explained by the increased monitoring in this group. The opposite was the case for contrast-enhancing lesions, where these occurred in a higher proportion of the untreated group [5/11 (45%)] than the early treated group [11/30 (37%)] despite differences in scanning frequency, but this was also not statistically significant. Drawing conclusions from these data is impossible owing to the higher frequency of scans undertaken in the early treatment group over a longer period of time but the fact there is no significant difference between cohorts despite this may favour the early treatment group but clearly this cannot be concluded from these data.

Safety

Given the increased use of DMTs in the early treatment group, more treatment-related AEs occurred in comparison to those who started treatment later. Most importantly, in the early treatment group, a patient died from natalizumab-associated PML. This occurred in a 35-year-old male diagnosed with Rapidly

Evolving Severe MS in 2011, aged 31, who was initially JCV antibody negative but seroconverted in late 2014 after approximately 3 years of natalizumab treatment. He had been relapse-free since natalizumab was started but presented in 2015 with a progressive brainstem syndrome and an enlarging pontine lesion on MRI brain. CSF was positive for JCV PCR and, despite plasma exchange and active management of his condition including ventilator support, he did not survive and was ultimately managed palliatively. No other SAEs were identified in the cohort. In terms of AEs leading to treatment discontinuation, this occurred in over half of DMT-treated patients in both groups, although was proportionally higher in the late treatment subgroup. Indeed, AEs were the most common reason for DMT discontinuation in both cohorts, more so than efficacy or disease progression as outlined in **Figure 2-13**. The early treated group had significantly more AEs per patient however (0.9 vs 0.3, p=0.014). These data are in keeping with the expectation that DMT-related AEs are the counterbalance to any efficacy benefits. However, only DMT-related AEs were identified as this was a retrospective analysis hence not all AEs would be recorded in routine clinical follow-up. This likely underestimates the incidence of AEs in both cohorts but, obviously, will only capture AEs in treated patients. MS-related disease activity and disability are the adverse outcome of no treatment and this has to be balanced with the risks of DMTs identified in pivotal trials and real-world follow-up such as this.

Figure 2-8: DMTs used years 1-6 in early treated cohort [n=39]

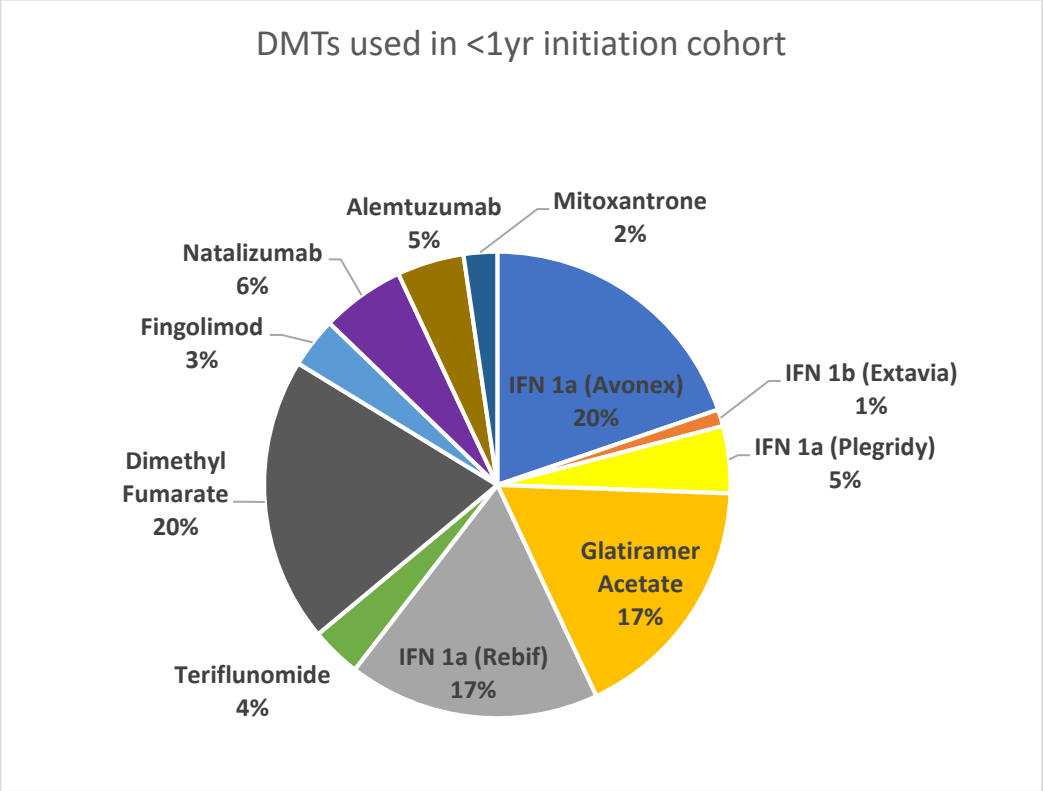


Figure 2-9: 1st DMTs used after 3 years in delayed initiation cohort [n=13]

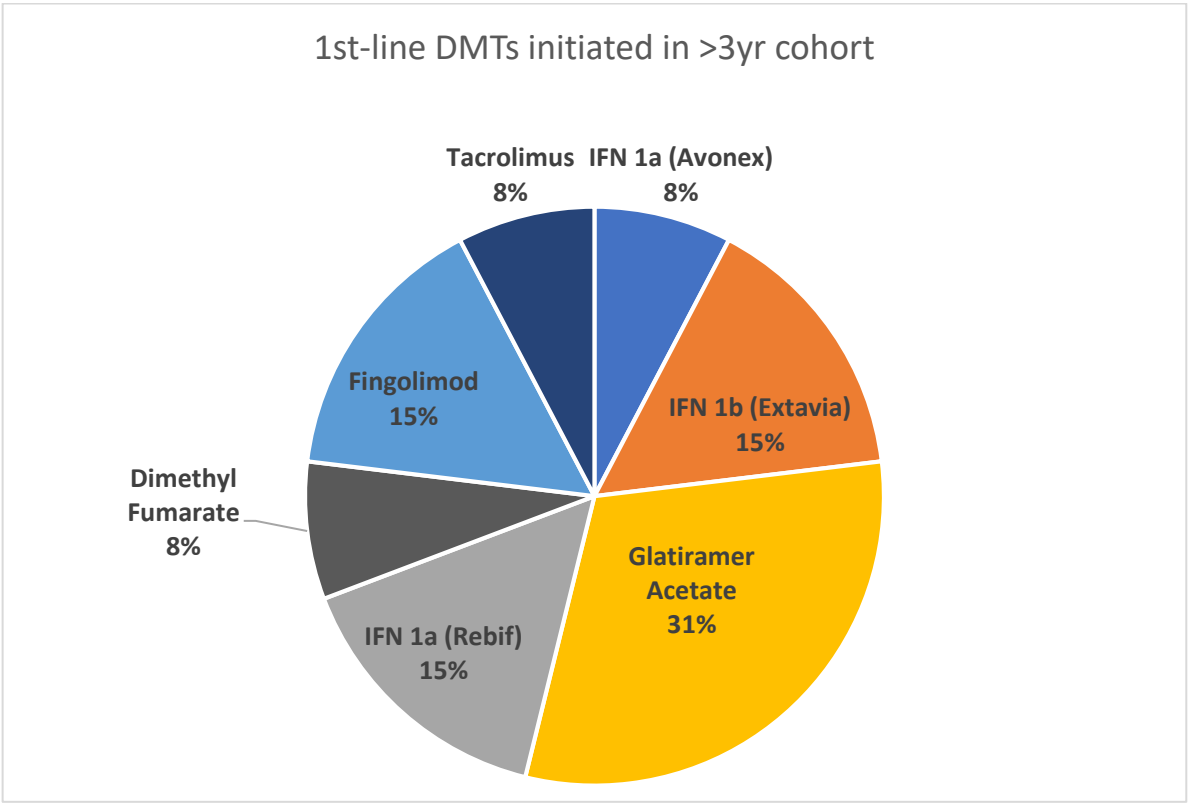


Figure 2-10: All DMTs used years 3-6 years in the initially untreated cohort [n=13]

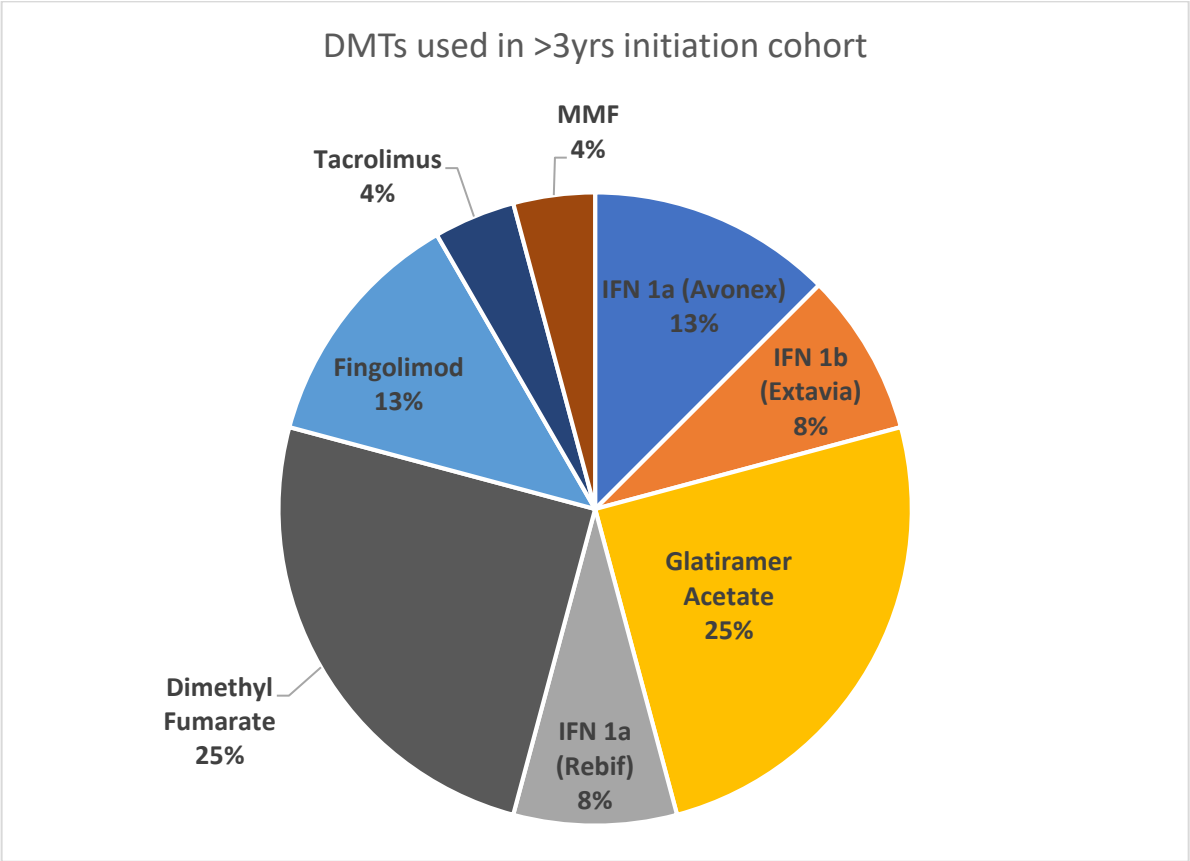


Figure 2-11: Comparison of DMTs used in early (<1yr) and late (>3yrs) initiation

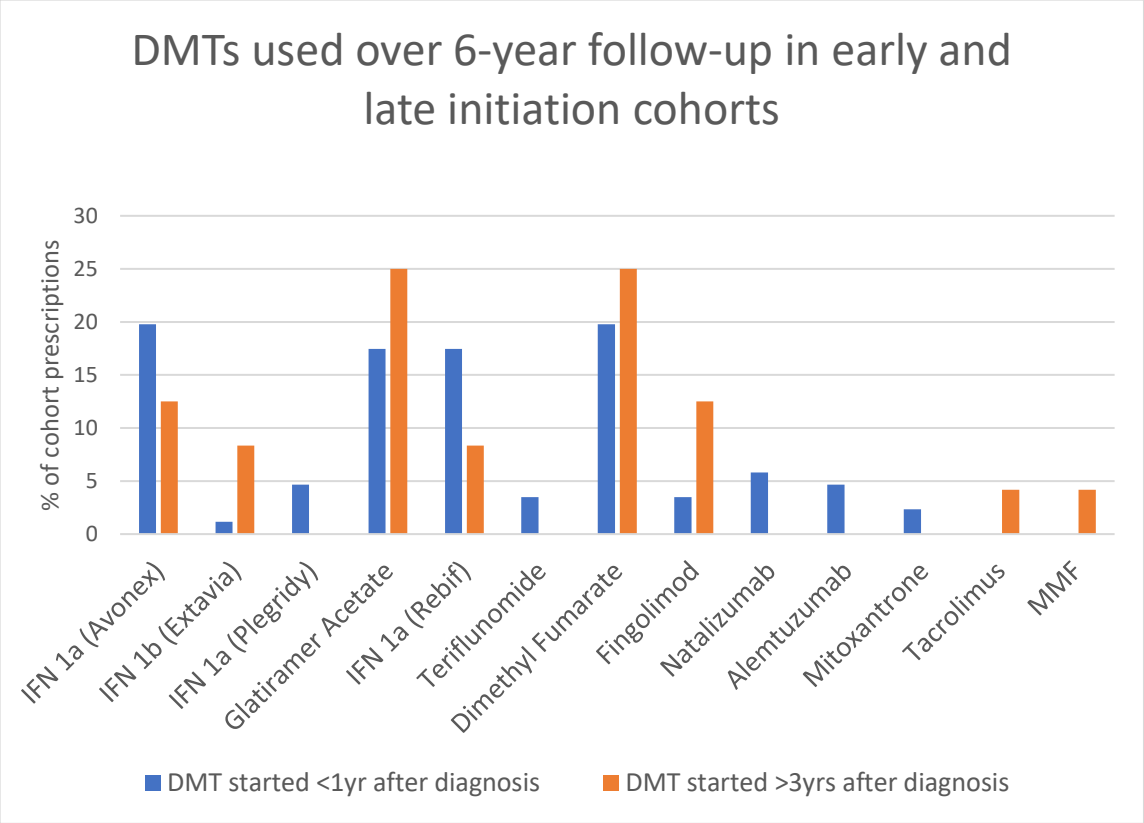


Figure 2-12: Symptomatic therapies used by each cohort

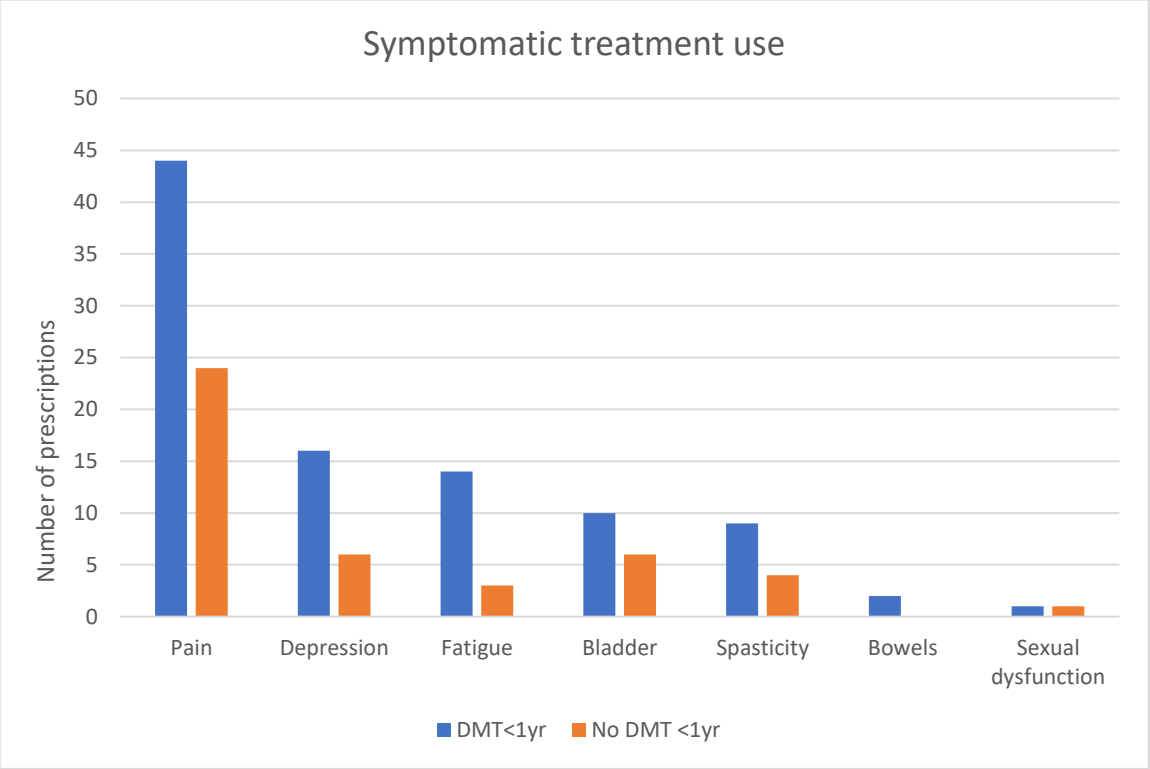
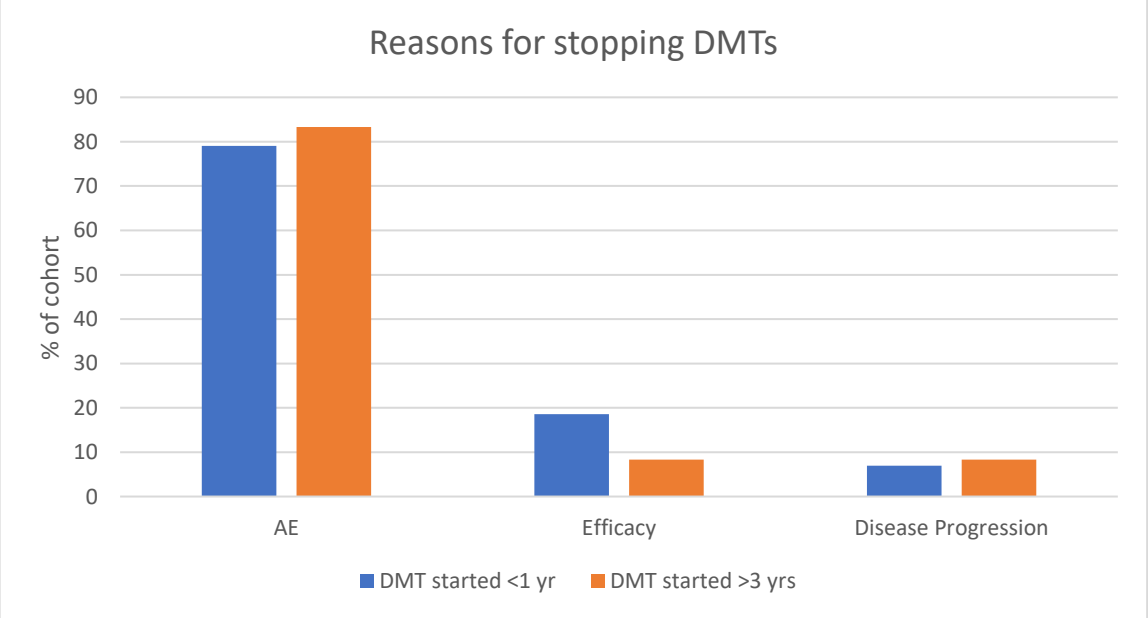


Figure 2-13: Reasons for stopping DMTs



Early vs late DMT initiation

- Early DMT initiation versus late DMT initiation
    - N = 36 (23 started DMT <1year vs 13 started DMT after 3 years, after relapse)
    - Groups 4,5 & 7 vs Group 3

In order to compare outcomes for early and late DMT initiation, we compared those who had started a DMT within 1 year from diagnosis and continued on treatment during follow-up but not escalated [Groups 4, 5 and 7 (N=23)] with those starting a DMT at least 3 years after diagnosis following relapse activity [Group 3 (N=13)]. **Table 2-11** outlines the comparisons between these cohorts.

There was no statistically significant difference between the cohorts at baseline although the early treated group had a higher absolute baseline EDSS score on average. The cohort sizes obviously differ however and the small numbers, particularly in the late treated group, reduce power to detect significant differences. Follow-up frequency and duration is much more comparable though as untreated patients have been excluded. Indeed, these cohorts have annual data available for around 5 years following diagnosis which resolves the issues in the entire cohort comparisons related to unequal follow-up and monitoring. Outcome comparisons between these cohorts would therefore be expected to more closely represent a true reflection of the intervention of early versus late DMT initiation. Of course, the earlier treated group had significantly longer exposure to DMTs than the later group (3.8 vs 1.7 yrs, p=0.002) and started treatments earlier following diagnosis (0.3 vs 3.6 yrs, p<0.0001) but there was no significant difference in the number of DMTs or symptomatic therapies used between the cohorts. **Figure 2-14** shows the DMTs used 1st-line in the early treatment group, with 1 patient receiving alemtuzumab (4%), 1 receiving mitoxantrone (4%) and 4 natalizumab (18%), hence around a quarter of this cohort used highly effective DMTs 1st-line and three-quarters injectable therapies. The vast majority of the early treated group were treated with a single DMT (17/23, 74%) but some patients switched to other therapies during follow-up, as outlined in **Figure 2-15**.

Relapses

The earlier treated group had significantly fewer relapses during follow-up in comparison to the later treated group. By definition, the later treated group were only escalated (from no treatment) following relapse however. This likely explains the absence of relapse-free patients in the later treated group but, additionally, ARR was significantly lower in the early treated group over time (0.23 vs 0.49, p=0.03). This remained the case when only relapses occurring at least 1 year after diagnosis were included in the analysis, in order to exclude immortal time bias. That said, as seen in the overall cohort not treated with DMTs within 1 year of diagnosis, the delayed treatment cohort had milder relapses and used steroids significantly less often for these.

Disability

There was no significant difference in disability outcomes between early and delayed DMT initiation groups. The frequency of EDSS monitoring was comparable between the groups, albeit the early treatment group had roughly annual documented scores whilst the later treated group had EDSSs available closer to a bi-annual basis but this was not significantly different. There was no significant difference in EDSS change from baseline to final EDSS, approximately 5 years apart on average for both groups, although the absolute numerical increase was higher in the later treated group (+1 vs +0.63, p=0.76). Notably, only 1 of the 13 later treated patients (8%) was considered to have developed SPMS during follow-up by the treating team, in comparison to 6 of the 23 early treated patients (26%) but the difference was not statistically significant. Similarly, a higher proportion of the early treatment group reached EDSS 3 which persisted until the end of follow-up but this was also not a significant difference.

MRI

The frequency of MRI monitoring was comparable between the early and later treated cohorts overall and no significant differences in disease activity were seen. A higher proportion of the early treated group had contrast-enhanced MRIs but there was no significant difference in the proportion developing new T2 or contrast-enhancing brain MRI lesions between the two cohorts.

Safety

DMT-related AEs were comparable between early and later treated patients but one patient in the early treatment group died from Natalizumab-associated PML, as discussed previously.

Table 2-11: Early vs late DMT initiation (N=36)

	DMT started < 1year and continued (Groups 4, 5 & 7)	DMT started after 3 years (Group 3)	p
Baseline characteristics			
N (% female)	23 (61)	13 (54)	
Mean age at diagnosis (SD)	36.8	35.6	
Mean no. Co-morbidities	1.2	1.8	0.255
Mean disease duration at diagnosis (SD) [years]	3.1 (n=22) (median 1.9)	2.8 (n=12) (median 2.3)	0.8 (t-test) 0.9 (Mann-Whitney)
Mean Baseline EDSS	3.5 (Median 3.5)	2.2 (Median 2.0)	0.050 (t-test) 0.052 (Mann-Whitney)
Baseline number of MRI brain T2 lesions (%) <9 >9 0	(n=19) 6 13 0	(n=12) 4 8 0	>1
Follow-up			
Mean duration of follow-up (yrs)	5.0	5.5	
Mean Number of visits per patient	7.6	7.3	
Mean duration from diagnosis to end of data collection	6.4	6.4	
Number of visits per year of study	1.2	1.2	
Mean number of visits per year of follow-up	1.9	1.3	0.24
Treatments			
Mean duration on any DMT (years)	3.8*	1.7	**0.002** (t-test)
Mean number of DMTs used	1.4	1.8	0.36
Mean duration from diagnosis to DMT commencement (years)	0.3*	3.6	**<0.0001** (t-test)
Mean number of symptomatic treatments used per patient	2.9	1.7	0.14
Relapses			

Total number of relapses recorded after diagnosis	32	32	
Proportion with 0 relapses after diagnosis	10/23* (43.5%)	0	**<0.0001 (z-test)**
Mean total number of relapses per patient after diagnosis	1.4	2.5	0.08 (t-test)
Mean duration to 1st relapse after diagnosis (years)	1.5 (n=13)	1.4	0.9 but missing data
Mean ARR during follow-up	0.23*	0.49	**p=0.03**
Mean ARR exc. Relapses <1yr after diagnosis	0.24*	0.49	
Severity of relapses (%)			
Mild	31.3*	65.6	**p=0.008 (z-tests)**
Moderate	46.9*	12.5	**p=0.003**
Severe	6.3	3.1	0.9
Unknown	15.6	18.8	
Steroids used	43.8*	11.5	**p=0.005**
EDSS			
Total EDSS scores recorded	95	45	
Mean EDSS scores documented per patient	4.2	3.5	0.45
Mean EDSS scores per patient per year of follow-up	1.0 (Median 0.8)	0.65 (Median 0.54)	0.22
Mean change from Baseline EDSS to final EDSS	+0.63 (Median 0) [n=19]	+1.0 (Median +0.25) [n=10]	0.76
Mean duration between baseline and final EDSS (years)	5.0 (Median 5.3) [n=19]	5.0 (Median 5.4) [n=10]	
Proportion developing SPMS	6/23 (26%)	1/13 (8%)	0.29
Proportion reaching EDSS 3 (maintained until last available EDSS) [denominator is number with ≥2 EDSS scores]	11/19 (58%)	4/10 (40%)	0.59
MRI			

Total number of MRIs undertaken [inc contrast]	84 [46] (n=19)	36 [9] (n=12)	
Number of patients having ≥1 MRIs with contrast	16	7	
Mean number of MRIs per patient	3.7	2.8	0.36
Mean MRIs per patient per year of follow-up	0.9 (Median 0.6)	0.5 (Median 0.6)	0.16
Patients developing new or enlarging T2 lesions	13/19 (68%)	7/12 (58%)	
Patients developing new Gd lesions	7/16 (44%)	3/7 (43%)	
Treatment-related adverse events			
Number of patients having AE leading to Rx discontinuation	7	6	
DMT-related AEs - total	10	9	
Mean number of AEs leading to Rx discontinuation per patient	0.6	0.7	
SAEs	Death from PML		

*p<0.05

Figure 2-14: First-line DMTs used in Early treatment cohort (Groups 4,5 & 7)

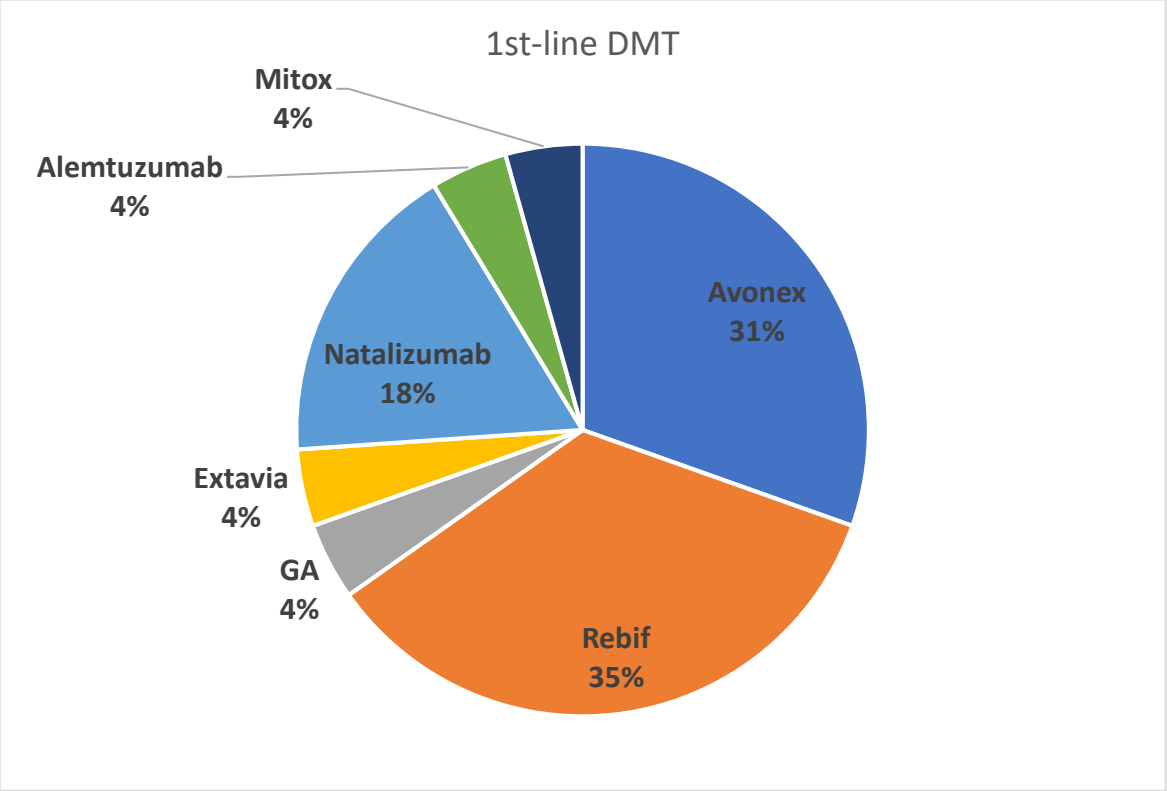
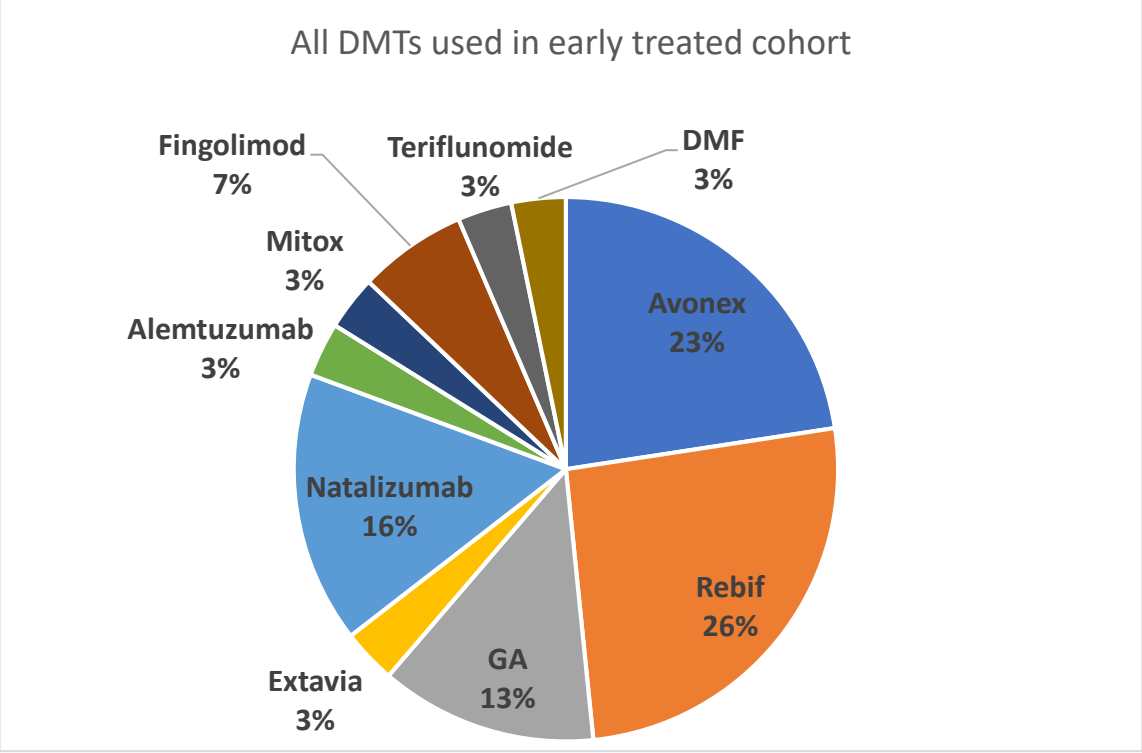


Figure 2-15: All DMTs used in early treated cohort (Groups 4,5 & 7)



Early DMT initiation (+/- escalation) vs. No DMT

Overall, there was no difference between the early DMT initiation with escalation cohort and the untreated cohort which was likely explained by the lack of follow-up of untreated patients, making longer-term comparisons impossible. **Table 2-12** outlines the comparisons between the two groups and highlights the significantly shorter follow-up of untreated patients than those receiving treatment (mean 2.2 vs 5.2 years, $p<0.0001$). Consequently, the number of clinic visits over the 6 years of potential follow-up was also significantly lower in the untreated cohort. The DMTs used in the treated cohort have been described previously and this group also used more symptomatic treatments on average than those not receiving a DMT although, again, this may simply be due to reduced access given the lack of follow-up.

Given these methodological issues, interpretation of clinicoradiological outcome comparisons cannot be conclusive. Longitudinal relapse outcomes clearly cannot be compared. The treated group had lower ARR than the untreated group but this was not statistically significant (0.3 vs 0.8, $p=0.052$) and disappeared when only relapses occurring at least 1 year after diagnosis were included. Conversely, but similar to other results, DMT-treated patients generally had more severe relapses but only the higher use of steroids was statistically significant. Baseline EDSS was notably higher in the early treated group but there was no significant difference in disability outcomes between the groups. Again, the treated group had proportionally higher rates of SPMS development and reaching EDSS 3 but this was not statistically significant. Similarly, there were no significant differences in MRI activity but the treated group had more recognition of new lesions likely as a result of the higher use of imaging in this cohort. The treatment-related AEs in the treated group have been described previously and clearly the untreated group was not subject to these.

Table 2-12: Early DMT and escalation vs no DMT (N=65)

	DMT started <1yr (Groups 4-8)	No DMT at last review / ever (Groups 1 & 2)	p
Baseline characteristics			
N (% female)	39 (62)	26 (81)	
Mean age at diagnosis (SD)	36.9 (8.7)	37.9	
Mean no. Co-morbidities	1.33	1.27	
Mean disease duration at diagnosis (SD) [years]	3.4 (4.4)	4.4 (n=22)	0.5
Mean Baseline EDSS	2.7* (Median 2.5)	1.4 (Median 1)	**p=0.009**
Follow-up			
Mean duration of follow-up (yrs)	5.2*	2.2	**<0.0001** (t-test)
Mean Number of visits per patient	7.7*	3.0	**<0.0001**
Mean duration from diagnosis to end of data collection	6.3	6.3	
Number of visits per year of study	1.2*	0.5	**<0.0001**
Mean number of visits per year of follow-up	1.7	2.6	0.13
Treatments			
Mean duration on any DMT (years)	4.2		
Mean number of DMTs used	2.1		
Mean duration from diagnosis to DMT commencement (years)	0.3 (median 0.3)		
Total number of symptomatic treatments used	96	18	
Mean number of symptomatic treatments used per patient	2.5*	0.7	**p=0.001**
Relapses			
Total number of relapses recorded after diagnosis	62	19	
Proportion with 0 relapses after diagnosis	12/39 (31%)	13/26 (50%)	0.19 (z-test)

Mean total number of relapses per patient after diagnosis	1.6*	0.7	**p=0.014 (t-test) p=0.02 (Mann-Whitney)**
Mean duration to 1st relapse after diagnosis (years)	1.8 (n=27)	1.6 (n=13)	0.83 but missing data
Mean ARR during follow-up	0.3	0.8	0.052
Mean ARR exc relapses <1yr after diagnosis	0.3	0.3 (n=14)	
Severity of relapses (%)			
Mild	34.9*	63.2	0.049
Moderate	38.7	15.8	0.06
Severe	6.6	0	0.3
Unknown	19.8	21.0	
Steroids used	35.5*	10.5	**p=0.02** (z-test)
EDSS			
Total EDSS scores recorded	158 (n=37)	44 (n=24)	
EDSS scores documented per patient	4.3*	1.7	**<0.0001**
Mean EDSS scores per patient per year of follow-up	0.9* (Median 0.8)	2.6 (Median 0.6)	**0.043** (t-test) 0.719 (Mann-Whitney
Mean change from Baseline EDSS to final EDSS	+1.1	+0.75 (Median +0.25) [n=12]	0.7
Mean duration between baseline and final EDSS (years)	5.2* (Median 6.0)	4.1 (Median 2.2) [n=12]	**p=0.013** (Mann-Whitney) **p=0.009** (t-test)
Proportion developing SPMS	8/39 (21%)	1/26 (4%)	0.32
Proportion reaching EDSS 3 (maintained until last available EDSS) [denominator is number with ≥2 EDSS scores]	17/33 (52%)	5/12 (42%)	0.8
MRI			
Total number of MRIs undertaken [inc contrast]	141 [79] (n=34)	48 [12] (n=24)	

Number of patients having ≥1 MRIs with contrast	30	6	
Mean number of MRIs per patient	3.6*	1.8	**p=0.003** (t-test)
Mean MRIs per patient per year of follow-up	0.8 (Median 0.6)	1.9 (Median 0.9)	p=0.123
Baseline number of T2 lesions (%) <9 >9 0	(n=34) 31.3 65.6 3.1	(n=24) 45.8 50.0 4.2	>0.05
Patients developing new or enlarging T2 lesions	19/34	7/24	0.066
Patients developing new Gd lesions	11/30 (37%)	1/6	0.57
Treatment-related adverse events			
Number of patients having AE leading to Rx discontinuation	21 (53.8%)		
DMT-related AEs - total	34		
Mean number of AEs leading to Rx discontinuation per patient	0.9		
SAEs	Death from PML		

*p<0.05

Discussion

This observational study has identified a cohort of patients with RRMS in Scotland with comparable inflammatory disease activity at baseline who were or were not treated with a DMT within 1 year of diagnosis. Overall, of 245 patients diagnosed with RRMS in Greater Glasgow and Grampian between 2010 and 2011, 130 patients (53%) were initiated onto a DMT within 1 year of diagnosis. PSM was used to generate pairs of patients from the treated and untreated cohorts with comparable disease activity at diagnosis based on established prognostic factors in order to identify treatment differences. In the unmatched cohorts, patients with poorer prognostic factors were significantly more likely to be treated. However, of the 225 patients with data available to calculate a PS, 124 (55%) were equally treated or not despite comparable PSs, suggesting factors other

than inflammatory disease activity and known clinical prognostic markers were driving the decision-making process. In this cohort of 62 matched pairs, it appears 'non-disease' patient and clinician factors are relevant to the decision to start a DMT. Notably, however, baseline EDSS was not a covariate in the PSM process and, when matched treated and untreated patients were evaluated for long-term outcomes, the treated group had a higher mean baseline EDSS. Therefore, although this was not evaluated in all 225 patients, this may be an additional disease factor which contributed to treatment decisions for both patients and doctors. Outcome comparisons between the treated and untreated cohorts were subject to detection bias as the treated patients had significantly more monitoring and longer follow-up. However, this was not an issue in comparing treated patients either staring DMTs early (<1 year after diagnosis) or later (>3 years), with over 5 years of follow-up available on average for both cohorts. The earlier treated group had significantly fewer relapses and lower ARR, although the later treated group had milder relapses when they occurred and there was no difference in imaging or disability outcomes.

Treatment variation in RRMS appears widespread globally but this has not been studied directly in the UK before. We used a survey of patients with MS in order to estimate DMT use in Scotland as there is no other systematic evidence of DMT prescription practices or access available[190]. Within Europe there are differences in the timing of DMT initiation and escalation but much of this can be explained by national prescribing guidelines[198]: there is a notable difference between treatment guidelines even within the UK, for example, with NHS England having a generally proscriptive approach. Prescribing differences within a single country, operating under the same national guidance, must be driven by different factors but also exists. In Sweden, for example, regional variations in healthcare, including the use of DMTs in RRMS, are evaluated regularly by the central government. In 2012, 46.6% of patients with RRMS in Sweden were treated with DMTs on average, but this varied from 25.8% in Varmland in the South to 79.4% in the Northern county of Vasterbotten[199]. Establishing treatment variation beyond regional differences, for single centres or prescribers for example, is obviously more challenging and has not been evaluated systematically. However, it is clear that DMT prescription practices in RRMS can vary for a number of potential reasons.

The reasons for discrepancies in the use of DMTs in similarly-affected patients with RRMS will be a summation of both clinician and patient factors. As we have shown, patients with more active disease will tend to use DMTs but the interaction between clinician and patient beliefs as to their relative benefits and costs likely results in varied use otherwise. In structured interviews with healthcare professionals involved in MS in the UK, a number of themes were identified which influence prescribing decisions[200]. Within the UK there are differing national guidelines which explain some of the variation between countries but other factors appear equally important. Variations in the definition of relapses impact upon the use of DMTs which are prescribed based on relapse frequency. The distinction between a true relapse and pseudorelapse can be difficult and severity is open to interpretation. Perceptions of the risks and benefits of DMTs in the mind of a treating physician and their familiarity with the drug in question was also highlighted as a factor in DMT utilisation in RRMS. Additionally, as most UK Neurologists don't work in isolation, the 'culture' and prescribing practice within their peer network also influences their individual decisions in discussions with patients. These two factors, prescribing experience and local culture, are established pivotal factors in variations in medical practice generally, summarised as 'geography and specialty are destiny' by some[201].

The discussion with the patient is also likely key to their ultimate decision on using DMTs or not. Very few patients wish to entirely devolve decision-making to their doctor, in a paternalistic model, and most appear to prefer a shared-decision model where patient and doctor agree on the strategy after considering the options[202–204]. The level of involvement in decision-making is variable amongst patients, however, with a passive role often preferred by older patients and those with lower educational attainment. Additionally, the level of risk associated with DMTs which patients will accept for any purported benefit is higher in males and those with established disability. However, the preferences and goals of patients may be at odds with their doctor[203,205,206] hence the final decision made will simply reflect that of the party with most influence if a truly shared-decision model is not used. Improved understanding and engagement of patients in treatment decisions is not only preferred by patients generally but also improves concordance. Risk-adversity is typically higher in clinicians than

patients, but disease-specialists and patients tend to overestimate the benefits of treatment and underestimate side effects which may reflect the reporting of misleading statistics in published trials[201]. Ultimately, however, it should be the case that patients have access to treatments for which they fulfil prescribing criteria and their doctors facilitate discussion as to the relative risks and benefits pertinent to the individual as far as possible before agreeing on a management strategy. The relative influence of these patient and clinician factors in our cohorts is not known, but recognising them is important if access to treatments is to be equitable.

Our study was limited in drawing conclusions about the impact of early DMT initiation on relevant clinical outcomes, but a recent study has attempted to answer this question using a similar approach but a much larger dataset[207]. Using the MSBase database to identify matched treated and untreated cohorts of patients with RRMS, Brown et al. also used PSM to compare treated and untreated patients, but with conversion to SPMS as the primary outcome. A standardised definition of SPMS was used, based on EDSS progression, which the group had previously published as highly correlating with irreversible subsequent disability worsening and identified half of a contemporary cohort of over 17000 patients as developing SPMS within 32 years from disease onset[100]. Similar to our study, treatment versus no treatment and early vs late treatment cohorts were compared. Whilst the treated patients were derived from multiple global centres (N=3960), the untreated control cohort was from a single UK centre (N=275) but all had at least 4 years of follow-up data. Also, in contrast to our study, PSM included baseline EDSS as a covariate and multiple matching (up to 10:1) was permitted such that a single patient could be matched multiple times in the analyses. Additionally, follow-up was censored to the shortest of the two follow-up times in matched patients, meaning an identical follow-up duration for each cohort. Overall, treatment with any DMT led to a significantly lower proportion developing SPMS in comparison to no treatment, with the greatest benefit using high efficacy treatment from the outset. For example, 11 years after diagnosis 47% of the IFN/GA-treated group had developed SPMS in comparison to 57% of the untreated group; at 8 years, 21% of the alemtuzumab-treated group had converted versus 41% of the untreated group and 7% versus 39% at 6 years in natalizumab-treated patients versus no treatment. Starting treatment earlier

(within 5 years of onset) rather than later was also beneficial, with 29% of the earlier treated group developing SPMS at 17 years in comparison to 47% of the later treated group. There was no significant difference in conversion rates to SPMS between untreated patients and those starting a DMT after 5 years when followed-up beyond 8 years however. Those who escalated from 1st- to 2nd-line DMTs in less than 5 years did benefit though, in comparison to later escalation, with 14% and 28% developing SPMS respectively after 7 years. The authors comment on the limitations of this study, however, which we could potentially help address in the prospective phase of MODERATE. Using our cohorts, we would have patients from a geographically comparable location, treated in the same era and ensure reliable clinical outcome measures undertaken specifically for the purpose of the study. Additionally, we would hope to include more extensive phenotyping of disease including cognitive and radiological outcome measures. We have also attempted to evaluate the risks of DMTs in our study, which was not addressed in the MSBase cohort.

Indeed, our current study has a number of strengths but its value lies mainly as a precursor to detailed evaluation of the treated and untreated cohorts in a prospective phase. The nature of the cohorts, diagnosed at the same time in a single country and healthcare system with identical prescription frameworks, excludes many confounding factors. A limitation of PSM is the reliance on known prognostic factors for inclusion in a treatment group, meaning unknown confounders may bias results in contrast to true randomisation, but our study minimises the chance of much of these. In observational studies, it is suggested that PSM is advantageous over other statistical models and caliper matching is robust when the sample size is small[208,209]. Additionally, with patients diagnosed in 2010-2011, we have significant duration of follow-up retrospectively for treated patients and can extend this to include all patients in a prospective study potentially.

Our study has a number of limitations, however, limiting any conclusions as to the risks and benefits of early DMT initiation from these results alone. From the outset, patients were drawn from only two centres and exclusions occurred due to lack of available medical records and missing data at all stages, given the reliance of routinely collected data accessed retrospectively. This had less of an impact on deriving the matched patients but highly influences conclusions

regarding clinical outcomes. The major issue was lack of follow-up of untreated patients in comparison to treated patients. Additionally, numbers are relatively small, although were robust enough to provide clinically and statistically significant results such as the lower rate of relapses in earlier treated patients. The lack of statistical significance in other clinicoradiological outcome measures likely belies the cohort size as much as any true difference between the cohorts. The flaws of PSM are well established, not least the possibility of unmeasured confounders[210-212]. Additionally, although sometimes considered a method of 'pseudorandomisation' the process is retrospective here and the choice of treatment was actually made by the patient and their treating physician and not by chance, which limits interpretation of the true effects of a treatment intervention in isolation. Patients in our study also did not have access to first-line oral DMTs at initiation, which differs from current practice, and diagnostic criteria have since been updated. Additionally, for outcome measures at least, our model did not balance for baseline EDSS between the cohorts and this is a known factor in disability over time. In any future prospective study, it would be critical to ensure the treated and untreated cohorts are balanced as far as possible on all known measures which may influence outcome and therefore this should be included as a covariate for PSM. The role of RCTs is not intended to be replaced by observational studies using PSM but is simply an attempt to provide some level of evidence, rather than none, where more robust trial designs are not available to answer clinically-relevant questions such as we have posed here.

Conclusions

We have identified treatment variation in patients with RRMS despite comparable disease characteristics in Scotland. Conclusions regarding the clinical impact of early DMT initiation were hampered by unequal follow-up and monitoring, with treated patients more closely followed, but evidence of benefit on relapse activity was seen in the early treated group. This is to be expected, in keeping with the results of short-term clinical trials, but the main focus of outcome evaluation in this study was long-term disability and no difference was found using the available retrospective data. Detailed evaluation of this and other clinicoradiological outcomes is planned in a prospective study of the patients identified in this pilot phase.

Chapter 3: Alemtuzumab after Natalizumab Switch in Evolving Rapidly Severe Multiple Sclerosis (ANSWERS MS)

Introduction

Natalizumab is a very effective monoclonal antibody approved in the UK & Ireland for use in rapidly evolving severe multiple sclerosis (RES MS), characterised by two or more disabling relapses in one year with signs of disease activity on MRI scanning. Although usually effective in this group of patients, a small proportion of patients fail natalizumab therapy, due to allergic or hypersensitivity reaction to the first few infusions, persistent disease activity (either clinical or radiological) on therapy, or adverse effects. Neutralising antibodies to natalizumab are often detected in patients hypersensitive to natalizumab, or who fail therapy due to persistent disease activity.

In clinical practice, natalizumab failure is a challenging dilemma. First-line disease modifying therapies (interferons and glatiramer acetate) are less effective than natalizumab and often patients who start natalizumab will have already failed such agents. Fingolimod is now an option for natalizumab failure in the UK, but its effectiveness in highly active disease is uncertain and based on limited data from clinical trials.

Alemtuzumab has been used ‘off label’ in the UK and Ireland for more than a decade in patients with highly active RRMS, both prior to and during the implementation of the pivotal phase 3 program (Comparison of Alemtuzumab and Rebif Efficacy (CARE) MS 1 and 2 trials). This ‘real-life’ clinical experience with alemtuzumab is unique to the UK and Ireland and relates to the fact that the drug was discovered and developed in Cambridge, where it has been used regularly as a treatment for RRMS since 1999. The CARE MS 1 and 2 studies have proven the effectiveness of alemtuzumab in RRMS, but long-term clinical experience outwith the UK and Ireland is limited to trialists and trial centres.

Both natalizumab and alemtuzumab are associated with significant safety concerns that limit their use, particularly in patients whose disease prognosis is potentially more favourable. Any additive effect and the risk of sequencing such therapies one after the other are unknown.

Alemtuzumab is now widely approved as a treatment for adults with active RRMS, defined by clinical or MRI features, but its use is often restricted either to patients failing first-line therapies or to patients with features of disease activity broadly similar to those currently characterised by the approved indication for natalizumab. In the absence of direct trial comparisons, data on response to alemtuzumab in patients with RES MS deemed suitable (in everyday clinical practice) for natalizumab, which has now been used in over 160,000 patients worldwide, will be very helpful for practicing clinicians.

Background

There has been considerable progress in the development of new DMTs for RRMS in the last 2 decades, hence the options for patients and treating clinicians have expanded significantly. Despite the use of such medications there remains a cohort of patients whose disease worsens and evolves. DMTs are generally used sequentially on a first, second or third-line basis when previous treatments have failed, with the least efficacious but safest treatments used first and the most efficacious but higher-risk treatments used last. This policy is known as 'escalation therapy' and is generally adopted after detailed risk / benefit discussions with individual patients. The recent emergence of oral agents for RRMS[65,123,213,214] with superior outcomes compared to injectable first-line therapies is currently changing the landscape of MS management, but the concept of 'escalation' remains accepted practice in most centres.

The experience of patients and treating clinicians in the UK and Ireland is currently unique in terms of data relating to the long-term real-world efficacy and safety of alemtuzumab. The place of alemtuzumab in the current complex therapeutic landscape of RRMS is uncertain and is strongly influenced by real-world experience of its efficacy and risk profile. In the ANSWERS study, we collected and analysed data on patients from the UK and Ireland who have been treated with alemtuzumab after natalizumab failure to control rapidly evolving severe RRMS.

Alemtuzumab

Alemtuzumab was developed in 1983 in the labs of Cambridge pathology, hence its original name of *Campath*. It was the first therapeutic humanised monoclonal antibody to be made and is directed against CD52, a surface glycoprotein of unknown function expressed on all mature lymphocytes. Its administration results in profound B- and T-cell lymphopaenia, resulting from antibody-dependent cell-mediated lysis. It is given as an intravenous infusion of 12mg/day for 5 consecutive days in the first cycle and for 3 consecutive days in the second cycle 12 months later. Recovery of B- and T-lymphocytes to the lower limit of normal takes 8 months and 3 years respectively[215]. Originally developed to treat B-cell chronic lymphocytic leukaemia, it has also shown efficacy when used (off-licence) in a number of autoimmune conditions[216-218]. Its therapeutic effect in MS is postulated to result from the complex reorganisation of the immune system which follows homeostatic reconstitution of lymphocytes after depletion, rather than the depletion itself[219].

Three open-label, one Phase 2 and two Phase 3 trials have studied the use of alemtuzumab in MS since 1991[49,99,120,220-222]. Four of these have been co-ordinated from Cambridge, UK where all patients continue to be followed up within an extension study.

The first use of alemtuzumab in MS included 58 patients with relapsing-remitting and secondary progressive disease. This was an open-label study conducted between 1991 and 2002 with a mean follow-up of 29 months[99]. In patients with established secondary progressive disease (n=36), a significantly reduced relapse rate [0.7/yr to 0.001/yr ($p<0.001$)] and absence of new lesions on MRI was accompanied by progressively worsening disability of the group over the study period, as measured by the Expanded Disability Status Scale (EDSS). When used in patients with RRMS (n=22), in whom licensed treatment had failed or who had a high early relapse rate suggesting poor prognosis, the annualised relapse rate (ARR) improved by 94% in comparison to the year before treatment, and mean EDSS improved by 1.2 points 2 years after a five-day course of 20mg Alemtuzumab/day intravenously. Following this, the concept of a treatment 'window of opportunity' to affect disease progression was suggested to explain the difference in the outcomes between patients with early versus established disease, the presence of secondary axonal damage being the postulated

mechanism for the lack of effectiveness in patients with progressive disease. Two further small open-label studies[220,221] (n=39 and 45 respectively) have since been published and showed similar reductions in ARR (-92% and -94% respectively), but more modest improvements in EDSS (0.36 and 0.38 respectively), albeit with shorter follow-up (mean 21 and 24 months respectively). Hirst *et al*[220] included a majority of drug-naïve patients (82%), whereas Fox *et al*[221] included only patients who had relapsed despite interferon therapy.

The emergence of autoimmune conditions in patients treated with alemtuzumab was apparent during these studies, as had been previously reported[114], with almost a third of patients developing autoimmune conditions following treatment. This was predominantly thyroid-related autoimmunity (Graves' disease most commonly), but one patient in the original series became dialysis-dependent after developing Goodpasture's syndrome[99].

The CAMMS223 trial was a multicentre phase 2 rater-blinded study, published in 2008, which randomised 334 treatment-naïve patients with RRMS to receive treatment with alemtuzumab or interferon beta 1a (INF-β1a)[222]. This trial examined patients with active, early disease of less than 3 years duration, a baseline EDSS of ≤3 and at least two relapses in the previous 2 years. In terms of clinical response, this trial was successful in establishing alemtuzumab as superior to INF-β1a in the prevention of relapse (relapse rates 0.11/year vs. 0.36/year) and the development of sustained disability measured by mean EDSS after 36 months follow-up. After 36 months, 80% of patients in the alemtuzumab group were relapse-free in comparison to 52% in the interferon group. Concerningly, the trial had to be suspended after 3 years due to the emergence of immune thrombocytopenic purpura (ITP) in three patients in the alemtuzumab group, one of whom died due to intracranial haemorrhage. However, five-year follow-up data of those who participated before the suspension has shown continued clinically significant benefits in the Alemtuzumab group despite the administration of only 1 cycle of treatment in the vast majority[223].

The Phase 3 Comparison of Alemtuzumab and Rebif Efficacy in MS trials (CARE-MS1 and CARE-MS2) were rater-blinded active comparator trials assessing the

effectiveness of alemtuzumab against INF-β1a, published in 2012[49,120]. CARE-MS1 analysed 563 previously treatment-naïve patients, while CARE-MS2 included 798 patients who had relapsed despite standard treatments. Both trials were conducted over 2 years. Given the different cohorts, the mean duration of disease was different (2.1 years in CARE-MS1 and 4.5 years in CARE-MS2). In both studies the ARR was reduced significantly in comparison to the INF-β1a groups: patients were roughly half as likely to relapse on Alemtuzumab compared with INF-β1a over the 2 years (CARE-MS1 hazard ratio = 0.45, CARE-MS2 hazard ratio = 0.51). In CARE-MS2 there was an overall improvement in EDSS worsening in the alemtuzumab group in comparison to interferon, though this was not seen in CARE-MS1. It was postulated that the lack of expected disease worsening in the INF-β1a group (perhaps as a consequence of relatively mild RRMS) made the study underpowered to detect a difference in effect on EDSS worsening between the treatments.

In all the comparator studies, the occurrence of any infection, serious infection, malignancy and autoimmune disease was higher in the alemtuzumab groups in comparison to IFN-β1a. Over 90% of patients receiving alemtuzumab experienced an infusion reaction, but these were largely mild, and prophylactic use of methylprednisolone and antihistamines was instituted for management of this complication. In the CARE-MS trials, 17.4% of patients randomised to alemtuzumab developed autoimmune thyroid disease compared to 5.3% of those taking IFN-β1a, and 0.8% (8 patients) developed ITP, with one patient requiring splenectomy[49,120]. Recent data have suggested that rates of autoimmune thyroid disease may rise with longer term follow up[224]. Three patients died in the alemtuzumab groups during the study period in total: two patients were involved in fatal road traffic accidents and one suffered a fatal aspiration pneumonia following a brainstem relapse of their MS. In addition, one patient died from sepsis after the study. Herpetic infections occurred in 16% of patients in the alemtuzumab group in CARE-MS1, a rate that was unchanged despite the use of prophylactic antivirals in CARE-MS2. This compared with a 2.4% rate of herpetic infection in the IFN-β1a groups across both studies. In CARE-MS1, two patients required hospital admission for herpes zoster infections. In the alemtuzumab arms of both CARE-MS trials a patient from an endemic region developed pulmonary tuberculosis during the trial period. In CARE-MS1, five patients (0.8%)

in the alemtuzumab group developed malignancy (basal cell, thyroid, vulval and colon carcinomas) while 2 patients (0.9%) developed malignancy in the IFN-β1a group (basal cell carcinoma, acute myeloid leukaemia). Two patients (0.5%) in the alemtuzumab arm of CARE-MS2 developed thyroid papillary carcinoma with no malignancies in the IFN-β1a arm. Drug discontinuation due to side effects was consistently lower in the alemtuzumab groups in comparison to interferon (1-3.3% vs. 6-12.1%)[225].

Following the European licensing of alemtuzumab (Lemtrada®), a standardized Risk Management Plan has been developed to monitor for adverse events encountered in clinical trials[226]. Baseline and subsequent monthly full blood count (with differential), serum creatinine and urinalysis with microscopy are required to screen for ITP and nephropathies respectively. These should continue until 48 months after the last infusion. Similarly, baseline and 3-monthly thyroid function tests should be monitored for this period to screen for potential thyroid dysfunction. Patients must be advised to comply with periodic tests and report possible adverse event symptoms early.

Natalizumab

Natalizumab is a humanised monoclonal antibody (IgG4) directed against transmembrane receptors (integrins) and prevents diapedesis of activated leucocytes across the blood-brain barrier, which limits the inflammatory response within the CNS[227]. Very late antigen 4 (VLA-4) is a critical adhesion molecule that regulates the translocation of leucocytes from the peripheral circulation to sites of inflammation. In binding to the α4 chain of integrins, natalizumab prevents its association with β1 integrins to form VLA-4. In contrast to alemtuzumab, its pharmacodynamic action is short-lived as, upon discontinuation of treatment, VLA-4 returns to baseline levels within 4 months[227].

Natalizumab is currently licensed worldwide as a first or second-line therapy for active RRMS although, in practice, it is most often used in patients who fail to respond adequately to first-line treatments such as interferon-beta or glatiramer acetate. The AFFIRM study[68] compared natalizumab to placebo and showed a significant effect of natalizumab on relapse rate, progression and MRI lesions.

Similarly SENTINEL[228] showed that adding natalizumab therapy to INF-β1a resulted in better outcomes compared to the combination of natalizumab and placebo.

In the AFFIRM trial, 627 patients were randomised to natalizumab treatment and 315 to placebo. The mean ARR at 1 and 2 years were the primary clinical endpoints: this was 0.26 and 0.23 at 1 and 2 years respectively in the natalizumab group, in comparison to 0.81 and 0.73 in the placebo group ($p<0.001$). In SENTINEL, the mean ARR was 0.38 and 0.34 in the natalizumab plus IFN-β1a group (n=589) at 1 year and 2 years, in comparison to 0.82 and 0.75 in the placebo plus IFN-β1a group (n=582). The cumulative probability of sustained progression of disability at 2 years was 17% and 23% in the active arms of AFFIRM and SENTINEL respectively, in comparison to 29% in both the comparator groups ($p<0.001$ for AFFIRM, $p<0.05$ for SENTINEL). At 2 years, 72% of patients in the natalizumab arm of AFFIRM and 61% in the natalizumab plus IFN-β1a arm of SENTINEL were relapse-free; this was significantly higher than the comparator groups [46% and 37% respectively ($p<0.001$)]. Notably, post-hoc analysis of AFFIRM[229] showed that, when stratified by baseline disease activity, those with highly active disease seemed to gain particular benefit from natalizumab therapy, as the proportion of patients free of both clinical and radiological disease was similar (65 vs. 69.5%) whether the disease was highly active or not at baseline. Overall, relapses were reduced by 81% and disease progression was reduced by 53% in patients with highly active disease at baseline[230]. These observations have informed the decision of the National Institute of Clinical Health and Excellence (NICE) in the UK to license natalizumab for RES MS[231].

The most concerning adverse outcome related to natalizumab is the risk of progressive multifocal leukoencephalopathy (PML). In AFFIRM, the only adverse events that were significantly more frequent with natalizumab than with placebo were fatigue and allergic reactions. There were 5 cases (0.8%) of anaphylaxis in the natalizumab arm. In SENTINEL, anxiety, pharyngitis, sinus congestion and peripheral oedema occurred more frequently with natalizumab in comparison to IFN-β1a, and one case of PML emerged. JC virus is the human polyomavirus responsible for PML and leads to lysis of oligodendrocytes, resulting in rapid nerve demyelination[232]. The case of PML in SENTINEL occurred after 29 doses of natalizumab in a patient also using IFN-β1a, which led to the

recommendation that dual therapy should no longer be used. Unfortunately, cases of PML have continued to develop in patients on natalizumab monotherapy since it was licensed for use. As of 31st August 2017, there have been 749 cases of PML from over 174 000 natalizumab-treated patients, with mortality between 20-25% and the majority of symptomatic survivors having a poor functional outcome[233]. Patients treated for 2 years or more, with prior exposure to immunosuppressants and positive anti-JC virus antibodies are at the highest risk (11.1. cases per 1000 patients), while patients with none of these factors have the lowest risk (0.09 cases per 1000)[188].

The risk of PML is dictated by a number of predisposing factors including prior use of immunosuppressants, increasing length of natalizumab treatment and JC virus seropositivity and serum JC antibody titre. Using these parameters, risk mitigation strategies are now employed to identify and counsel patients regarding their risk of PML. To date, PML has not been reported in alemtuzumab-treated RRMS patients although did occur when alemtuzumab was used for haematological neoplasia[234].

These extensive clinical development programs have demonstrated that both natalizumab and alemtuzumab are extremely effective in controlling RRMS, but that both are limited by their risk of serious adverse events. The mechanism of action and dosing schedules for each drug is notably different, and the risk-benefit equation for individual patients is dictated by the severity and prognosis of their disease, the likely benefits of each drug, and the risks and toxicity associated with each drug.

Study Rationale

In the coming years, the current staged 'escalation' approach to early management of RRMS may be challenged if long-term follow-up studies strengthen the 'window of opportunity' concept and suggest that early effective DMT use can prevent or delay the development of persistent or progressive disability. There may be a move to an induction therapy approach using the most efficacious treatments early, particularly in patients with clinical or radiological evidence of active disease. With either strategy the decision about when, how and why to use each drug will evolve as experience of newer drugs grows in the

global MS community. The optimal sequencing of the vast array of treatments now available is unknown and various therapeutic paradigms have been suggested, including sequential monotherapy, induction and maintenance therapies, combination therapy or highly personalised individual treatment plans if an accurate prognosis can be predicted[235]. Multiple pharmacological (mechanism of action, pharmacodynamic effects), clinical (safety profiles, teratogenicity and pregnancy risk, method of administration, need for monitoring, overall efficacy) and non-clinical (cost, availability, licensing) factors will influence this, in addition to the individual assessment of benefit and risk that must be made by each patient.

The ANSWERS study provides real-life clinical data on the use of alemtuzumab after natalizumab failure in patients with rapidly evolving severe relapsing remitting multiple sclerosis. Head-to-head studies of disease modifying agents are unlikely to be conducted in the current climate, and observational studies of effectiveness in distinct clinical scenarios currently provide the best data for practising clinicians to make complex decisions about both the initial choice and sequence of MS disease modifying therapies.

Aims

Describe the real-world use of alemtuzumab after natalizumab in rapidly evolving severe RRMS and the safety and efficacy of this treatment sequence.

Primary objectives

1) To evaluate the safety and tolerability of using alemtuzumab after natalizumab failure

2) To describe the occurrence of relapses, the occurrence of sustained accumulation or reduction of disability, and the development of significant MRI markers of active inflammatory disease during each of five treatment phases (Pre-Natalizumab, Natalizumab treatment, Switch phase, Alemtuzumab treatment and Post-Alemtuzumab)

3) To assess the number of patients without relapse, significant MRI markers of active disease, or sustained accumulation of disability ('freedom from disease activity') in each phase

Secondary objectives

1) To evaluate whether there is an optimal time period for the 'switch' phase between natalizumab and alemtuzumab and the effect of alternative DMTs as bridging therapies

2) To describe any effects of natalizumab and alemtuzumab on qualitative atrophy and cognition in patients with rapidly evolving severe RRMS

Primary endpoints

1) The incidence of infections (including opportunistic infections) during each treatment phase
2) The incidence of secondary autoimmunity during each treatment phase
3) The incidence of infusion reactions, hypersensitivity reactions, new blood or urine monitoring events or neoplastic events during each treatment phase
4) Annualised relapse rate during each treatment phase
5) The occurrence of sustained accumulation of disability (defined by an increase from baseline of at least one EDSS point or ≥1.5 points if baseline EDSS score was 0) confirmed over 6 months in each treatment phase
6) The occurrence of significant MRI evidence of active inflammatory disease in each treatment phase
7) 'Freedom from disease activity' in each treatment phase

Secondary endpoints

1) The incidence of clinically significant or disabling relapses during each treatment phase

2) The occurrence of worsening or improving T2/FLAIR lesions, new T1 lesions and changes in qualitative atrophy during each treatment phase
3) The occurrence of new clinical markers of cognitive impairment during each treatment phase
4) Ambulation score in each treatment phase
5) EDSS change during each treatment phase
6) Presence of neutralizing antibodies in each treatment phase
7) Change in any marker of quality of life / employment

Study Population

All patients treated with alemtuzumab due to natalizumab failure in the UK and Ireland, out with clinical trials.

Methods

This was a retrospective observational study using routinely collected clinical and radiological data. Clinicians from 13 MS specialist centres in the UK & Ireland, identified as having suitable patients for the study through personal communication, entered data into a bespoke online database created by an independent software engineer and hosted securely by NHS Scotland.

Data Capture System

Software Development

An electronic database was created in conjunction with a software company (TechNeoSoftwareSolutions©) in order to capture a standardised dataset for each patient, from each centre. The following section, developed in conjunction with a software engineer (Mark Adamson) summarises the software methods used to develop the database.

The architecture used for the AnswersMS application is the Model-View-Controller (MVC).

MVC is a model used for application development; it separates the three main components of the application into its own isolated environment. These three components are the Controller, the View, and the Model

The Controller manages the requests, determining which view needs to be loaded, and interaction with the models. For example, when you go to login to AnswersMS, the Controller tells the application that the login View is to be loaded. When logging in, the controller loads the Model that handles logins, which will check the database that the credentials match. If the login attempt is successful, the Controller will redirect the user to the patient screen, the patient controller will then handle any further requests.

The View is the GUI that the user interacts with. The Controller will call the View after getting the required information from the Model. Basically, the View is what the user will see on the screen (the web page).

The Model is where data from the controller is actually passed and interacted with. The Model, will take the username and password passed from the Controller, verify that information against what is stored in the SQL database, then communicate with the Controller to display the correct View. For example, if an incorrect username or password is entered, the Model will pass this information to the Controller that the credentials are incorrect, then the Controller will communicate to the View to display an error message e.g. “Your username or password is incorrect.

The Technologies used to create Answers MS consisted of Microsoft SQL Server 2014 and Visual Studio 2013. To implement a work stream within the application the developer should start with creating a database table in SQL Server.

Upon creation of the patient table (**Appendix 1**) the developer now has a structure for implementing the application stream in Visual studio.

Following the MVC pattern the developer can select the database table to underpin the dataset used in the application.

Firstly, the developer is required to add a controller to manage request for the patient object (**Appendix 2**).

The model class is selected from the list e.g. patient table and the views for the user interface are created. MVC links all the required code together and allows the developer to focus on implementing the complex functionality on the system.

Once the developer has a complete work stream, they will repeat the process for each table they wish to interact with in the application.

User management is controlled by aspmemebership**.** ASP.NET Web Site Administration Tool can be accessed by clicking ASP.NET Configuration from the Website menu.

The developer can programmatically access the features provided by the ASP.NET Web Site administration tool via the System.Web.Security namespace in the application. Membership and Roles are used to store, access and modify user information in the application database. The user could be authenticated using the Membership.ValidateUser. Page-based user authorization is given by using the AuthorizeRequest event in the HttpApplication class.

Clinicoradiological data fields

The standardised dataset requested from each centre for each patient was based on the primary and secondary objectives and endpoints of the study as outlined above. These included:

- detailed descriptions of disease activity and severity (relapse rates, EDSS scores) during the 2 years prior to natalizumab therapy
- patient serial JC virus status (seropositive / seronegative / JCV titre)
- behaviour of patients on natalizumab (relapse rates, EDSS scores, presence of hypersensitivity / infusion reaction symptoms, presence of natalizumab antibodies)
- total natalizumab treatment course, as well as the reason for discontinuation
- interval between natalizumab and alemtuzumab
- infusion reactions to alemtuzumab

- disease activity during the years of follow up after alemtuzumab (relapse rates, EDSS scores)
- side effects and complications of alemtuzumab therapy (specifically lymphopenia, thrombocytopaenia and organ-specific autoimmune disease)
- presence of any opportunistic infections, including progressive multifocal leukoencephalopathy (PML), during the course of sequential 'aggressive' monoclonal antibody therapy
- presence of any other, perhaps previously unreported, complications that might be attributable to sequential monoclonal antibody therapy
- requirement for further alemtuzumab courses
- requirement for further disease modifying therapy / treatment for RRMS

Where formal EDSS scores were unavailable, these were estimated by the local MS specialist clinician using *Neurostatus Version 04/10.2* as a reference aid, which was incorporated into the online database, along with other prompts including the efficacy and safety definitions outlined below.

The clinical and radiological databases were planned and designed using Microsoft Powerpoint® as outlined in **Appendix 3 and 4**.

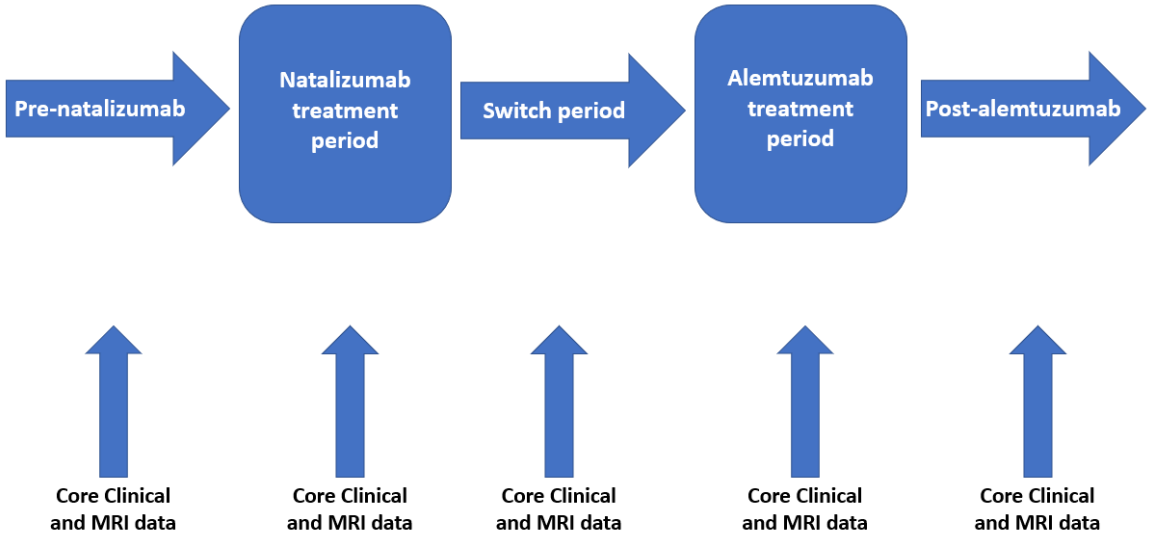
Longitudinal MRI imaging was obtained for all patients where available. A standardised MRI dataset was completed for each patient at each site by a blinded neuro-radiologist, using the online database. The (neuro)radiologist documented changes in MRI over time (see MRI data section below) using the available interval MRI scans that had been conducted for each patient. The effect of sequential natalizumab and alemtuzumab therapy on MS radiological lesion development and activity could thereby be determined in each patient, and summarised for the whole cohort.

Data were anonymised so that no patient identifiable information was provided within the online database or accessible by persons outwith the clinical team caring for each individual patient.

Study Design

Duration of follow up varied between patients and all data collection was of course subject to availability at each centre. Clinical data was collected for 6 phases as shown in Figure 3-1. Phase 1 (Baseline) was defined by the date of diagnosis, with data from that timepoint (or before) included only. Phase 2 (Pre-natalizumab) was defined as the period between diagnosis and the first dose of natalizumab. Phase 3 (Natalizumab treatment) was the period from first natalizumab infusion until 4 weeks after the last natalizumab infusion, to account for persisting pharmacodynamic effects. Phase 4 (Switch) was the time period between the last dose of natalizumab and the first dose of alemtuzumab. Phase 5 (Alemtuzumab treatment) was pre-defined as 2 years following the first infusion of alemtuzumab. Phase 6 (Post-alemtuzumab) was the time from the end of the alemtuzumab treatment phase until the end of data collection (31/12/17 for clinical data; 31/1/18 for radiological data).

Figure 3-1: Study Design



The clinicians and neuro-radiologists involved in data entry were provided with a centre-specific guide on the use of the online database. Patients with the longest follow-up since starting alemtuzumab were prioritised for data entry, specifically those with at least 2 years follow-up after alemtuzumab was commenced. Whilst centres accumulated new cases as the study progressed, with the majority of patients receiving alemtuzumab in the licensed era, those with at least 2 years follow-up after alemtuzumab were the focus of this study.

Clinical Data

The following data were collected for each patient in each phase:

1) Phase 1: Baseline

a) Patient-specific data*
   i. Demographic data
   ii. Co-morbidities
   iii. Deprivation category (based on postcode)
   iv. Employment status at diagnosis
   v. Pregnancy data

b) Para-clinical test results
   i. Cerebrospinal fluid results

ii. Visual evoked responses results

c) Disease-specific data
   i. Age /date at MS diagnosis
   ii. Presenting symptom(s) of MS
   iii. Date of first (likely) symptom if remote from presentation
   iv. EDSS / ambulation score at diagnosis
   v. MRI data**

d) Relapse information ***
   i. Number of relapses prior to diagnosis
   ii. Severity of relapses
   iii. Steroid use

2) Phase 2: Pre-natalizumab

a) Disease-specific data
   i. EDSS / ambulation score at 6 monthly intervals
   ii. MRI data**
   iii. Symptomatic therapy use

b) Relapse information ***

c) Other disease modifying therapy (DMT) use
   i. Date started and duration of first and any subsequent DMT
   ii. Adverse events/serious adverse events/side effects ¶
   iii. Presence of interferon neutralizing antibodies¶

d) JC virus status / index¶

3) Phase 3: Natalizumab Treatment

a) Date natalizumab started / Age at 1st dose
b) EDSS /ambulation score on day 1 of natalizumab therapy
c) EDSS /ambulation score at 6 monthly intervals on natalizumab
d) Relapse information***
e) MRI data**
f) Cognitive data***

g) Quality of life data***
h) JC virus status / index¶
i) Duration of Natalizumab treatment / number of infusions
j) Adverse events/serious adverse events/side effects¶
k) Reason for failure of Natalizumab
    i. Adverse events
    ii. Unsatisfactory efficacy
    iii. Risk mitigation
l) Presence of neutralising antibodies¶

4) Phase 4: Switch period

a) Age / date at start of switch phase
b) EDSS at 6 monthly intervals during the switch phase
c) Duration of switch phase
d) Relapse information***
e) MRI data**
f) JC virus status / index¶
g) Symptomatic therapy use
h) Cognitive data***
i) Quality of life data***
j) DMT use
    i. Date started and duration of any DMT
    ii. Adverse events/serious adverse events/side effects¶
    iii. Presence of interferon neutralizing antibodies¶

5) Phase 5: Alemtuzumab treatment period (2 years)

a) Date of Alemtuzumab treatment / Age at 1st infusion
b) EDSS / ambulation score at 6 monthly intervals after the first Alemtuzumab infusion
c) Relapse information***

d) MRI data**
e) Cognitive data***
f) Quality of life data ***
g) Symptomatic therapy use
h) Number of Alemtuzumab infusions / total Alemtuzumab dosage
i) Adverse events/serious adverse events/side effects¶

6) Phase 6: Post-alemtuzumab

a) Age / date at start of post-Alemtuzumab phase
b) EDSS /ambulation score at 6 monthly intervals during the post-Alemtuzumab phase
c) EDSS at the end of the post-Alemtuzumab follow up period (last score)
d) Number of further Alemtuzumab infusions/ total Alemtuzumab dosage
e) Relapse information***
f) Adverse events/serious adverse events/side effects¶
g) MRI data**
h) Cognitive data***
i) Quality of life data***
j) Symptomatic therapy use
a) Other disease modifying therapy (DMT) use
    i. Date started and duration of any DMT
    ii. Adverse events/serious adverse events/side effects¶
    iii. Presence of interferon neutralizing antibodies¶

*These data were collected for each phase and updated accordingly
**See MRI data section below
***See Efficacy Definitions below
¶See Safety Definitions below

MRI data

A neuroradiologist was identified in each centre by the centre lead.

The neuroradiologist was blinded to all clinical data and was provided only with the serial images for each patient. They were informed that the study was being conducted to assess the serial efficacy and safety of commonly used disease-modifying therapies in multiple sclerosis but were not informed which drugs or sequence of drugs were being assessed or which treatments patients were receiving at the time of each scan.

The neuroradiologist was provided with all available MRI brain images for each patient from diagnosis through the 6 assessment stages of the study (Baseline, Pre-natalizumab, Natalizumab treatment, Switch, Alemtuzumab treatment, Post-alemtuzumab). Every effort was made to obtain all films for each individual, but some were unavailable. Additionally, as a retrospective study, MRI scans were undertaken at different frequencies and intervals depending on clinician judgement at the time, likely related to disease activity and/or DMT use i.e. higher frequency of imaging in those with more clinically active disease or switching DMTs.

The first available scan for each patient was the Baseline scan. The blinded neuro-radiologist commented on supra- and infra-tentorial lesion load (minor, moderate, marked), the presence of significant atrophy (global / corpus callosum) and the number of gadolinium enhancing lesions on the baseline scan.

The following data were entered by the neuro-radiologist into the online database for each subsequent scan, making a comparison with the preceding scan:

1) Date scan was conducted
2) Number of new Gadolinium enhancing lesions
2) Number of new T2/FLAIR lesions
3) Improved (smaller or less intense) T2/FLAIR lesions
4) Worsened (larger or more intense) T2/FLAIR lesions
5) New T1 hypo-intense lesions

Specific MRI brain scans were also compared in a bid to assess the radiological efficacy of natalizumab and alemtuzumab. The scans to be compared were

identified automatically by the database, if available, once all available scan dates were entered by the neuro-radiologist. The database referenced the date of each scan with the dates of each treatment phase entered via the clinical database, to which the neuro-radiologist had no access. The specific MRI brain comparisons were:

1) Latest scan in Natalizumab treatment phase vs. latest scan in Pre-natalizumab phase [Effect of natalizumab treatment]
2) Latest scan in Alemtuzumab treatment phase vs. latest scan in Switch phase [Effect of alemtuzumab treatment within first 2 years]
3) Latest scan in Post-alemtuzumab phase vs. latest scan in Switch phase [Overall effect of alemtuzumab]
4) Latest scan in Post-alemtuzumab phase vs. latest scan in Alemtuzumab treatment phase [Effect of alemtuzumab in the longer term]

In these comparisons, the scan listed first is the more recent one, compared with (vs.) a previous one, with the following data collected:

1) Date scan was conducted
2) Number of new Gadolinium enhancing lesions
2) Number of new T2/FLAIR lesions
3) Improved (smaller or less intense) T2/FLAIR lesions
4) Worsened (larger or more intense) T2/FLAIR lesions
5) New T1 hypointense lesions
6) Qualitative Atrophy (Global and corpus callosum): improved, unchanged or worsened

The blinded neuroradiologists were additionally provided with free text areas within the database to make additional comments if necessary.

Serial MRI imaging of spinal cord was also retrieved where available, both cervical and thoracic specifically. Similar comparisons were made between baseline scans and all subsequent follow-up scans. For the baseline spinal cord (cervical/thoracic), the following data were entered:

1) Number of T2 lesions

2) Number of Gadolinium enhancing lesions
3) Qualitative atrophy assessment (Present or Not)

For subsequent interval spinal cord imaging, the number of new T2 and Gadolinium enhancing lesions were documented and whether atrophy had changed ('improved', 'unchanged' or 'worsened') in the opinion of the neuroradiologist.

Efficacy Definitions

These definitions were embedded within the online database as a reference for clinicians entering data.

Relapse

A relapse was defined as the onset of new symptoms or the worsening of pre-existing symptoms attributable to demyelinating disease lasting for more than 24 hours and preceded by improving or stable neurological status for at least 30 days from the onset of the previous relapse in the absence of infection, fever or significant metabolic disturbance. Objective change on neurological examination was not necessary to fulfill relapse criteria – the occurrence of relapse was ultimately up to the investigators' judgement.

Relapse severity was judged by a relapse severity score, dividing relapses into the following categories:

a) Clinically Significant Relapse - a relapse lasting more than 48 hours, fulfilling one of the criteria below:

    - Any motor relapse
    - Any brainstem relapse
    - A sensory relapse if it leads to functional impairment
    - Optic neuritis
    - Intrusive pain

b) Disabling relapse - any relapse which fulfilled one or more of the following criteria:

- Affects the patient's ability to work
- Affects the patient's activities of daily living as assessed by an appropriate method
- Affects motor or sensory function sufficiently to impair the capacity or reserve to care for themselves or others as assessed by an appropriate method
- Needs treatment (steroids)/hospital admission

The need for steroid therapy was also recorded.

EDSS score

EDSS scores were variably documented in clinical notes, but all centre leads and investigators involved in the project are very familiar with the EDSS score, its scope and its limitations. EDSS scores were entered whenever formally stated in clinical notes. Additionally, EDSS scores were estimated by clinicians entering data if a clinical examination (but not formal EDSS) was documented upon which to base this (using *Neurostatus* EDSS scoring sheet version 04/10.2 as a guide) at the following time-points (or more if available):

a) at baseline (date of diagnosis)
b) at 6 monthly intervals during the Pre-natalizumab period
c) at 6 monthly intervals during natalizumab treatment
d) at 6 monthly intervals during the Switch phase
e) at 6 monthly intervals after the first Alemtuzumab infusion
f) at 6 monthly intervals during the post-Alemtuzumab phase
g) at the end of the post-Alemtuzumab follow up period (last score)

Ambulation score was also estimated at each of the above time-points if appropriate information was available.

Sustained accumulation of disability (SAD) was defined as an increase in EDSS, sustained for at least 6 months, of ≥1.5 EDSS points if the baseline EDSS was 0;

≥1.0 point if the baseline EDSS was ≥1 but <5.5; and >0.5 points if the baseline EDSS was ≥5.5.

Sustained reduction in disability (SRD) was defined as a reduction in the EDSS score of either ≥1.0 or 0.5, for baseline EDSS scores below and above 5.5 respectively sustained for at least 6 months. Analysis for SRD was restricted to those with a baseline EDSS ≥2.0.

MRI assessment definitions

Significant MRI evidence of active inflammatory disease was defined as 1 new gadolinium
enhancing lesion and / or 2 or more new T2/FLAIR lesions on an interval scan compared to the previous scan.

Cognition assessment

Clinicians were asked to provide the results of any of the following cognitive assessments which were documented during each phase:

- Mini Mental State Examination (MMSE)
- Addenbrooke's Cognitive Examination (ACE)
- Montreal Cognitive Assessment (MOCA)
- Paced Auditory Serial Assessment Tool (PASAT)

Quality of life assessment

Formal quality of life measures were not made, but data on employment status, driving and changes in deprivation category (based on Postcode) were collected.

Safety Definitions

Adverse event (AE)

Any untoward medical occurrence that did not necessarily have a causal relationship with natalizumab, alemtuzumab or any other DMT that has been used. An AE was therefore any unfavourable and unintended sign (including an

abnormal laboratory finding), symptom, or disease occurring during the five treatment phases, whether or not related to any medicinal product.

Adverse events were grouped into:

a) Infusion reaction – any adverse event occurring during or within 24 hours of natalizumab or alemtuzumab infusions
b) Hypersensitivity reaction – reports of hypersensitivity, allergic reaction, or anaphylactic or anaphylactoid reaction by the investigator, as well as any report or urticarial, allergic dermatitis or hives
c) Blood monitoring event – occurrence of renal, liver or thyroid function test abnormalities, leucopaenia, lymphopenia, neutropenia, anaemia or thrombocytopenia
d) Urine monitoring event – the presence of proteinuria, haematuria, red cell casts or abnormal urine microscopy
e) Organ specific autoimmune disease – the occurrence of any organ specific autoimmune disease, specifically thyroid disease (including its nature/extent) and immune thrombocytopenic purpura
f) Neoplastic event – the occurrence of any neoplasm
g) Infection event (including opportunistic infection) – any infection event, specifically any opportunistic infection, any herpetic infection (including shingles), any viral infection or any bacterial infection, whether confirmed by diagnostic studies or not

Serious Adverse Event (SAE)

An SAE was defined as any untoward medical occurrence that at any dose:

a) Resulted in death
b) In the view of the centre lead, placed the subject at immediate risk of death (a life-threatening event); however, this does not include an event that, had it occurred in a more severe form, might have caused death
c) Required inpatient hospitalization or prolongation of existing hospitalization
d) Resulted in persistent or significant disability/incapacity, or

e) Resulted in a congenital anomaly/birth defect

An SAE could also be any other medically important event that, in the opinion of the centre lead, jeopardised the subject or required intervention to prevent one of the other outcomes listed in the definition above. (Examples of such medical events include allergic bronchospasm requiring intensive treatment in an emergency room or convulsions occurring at home that do not require an inpatient hospitalization.)

Again, serious adverse events were grouped into:

a) Infusion reaction – any serious adverse event occurring during or within 24 hours of natalizumab or alemtuzumab infusions
b) Hypersensitivity reaction – reports of severe / serious hypersensitivity, allergic reaction, or anaphylactic or anaphylactoid reaction
c) Blood monitoring event – occurrence of severe or significant renal, liver or thyroid function test abnormalities, leucopaenia, lymphopenia, neutropenia, anaemia or thrombocytopenia
d) Urine monitoring event – the presence of significant or marked proteinuria or haematuria
e) Organ specific autoimmune disease – the occurrence of any severe or significant organ specific autoimmune disease, specifically resistant thyroid disease, immune thrombocytopenic purpura, Goodpasture's syndrome and glomerulonephritis
f) Neoplastic event – the occurrence of any neoplasm
g) Infection event (including opportunistic infection) – any serious infection event, specifically any severe opportunistic infection, any evidence of PML (progressive multifocal leukoencephalopathy), any severe herpetic infection, any serious viral infection or bacterial infection, whether confirmed by diagnostic studies or not.

<u>Neutralising antibodies</u>

Natalizumab neutralising antibodies generally occur in 5-10% of patients receiving natalizumab and are strongly associated with both infusion-related

adverse events (including hypersensitivity) and loss of efficacy. The presence of neutralising antibodies was documented where undertaken.

The occurrence of natalizumab antibodies was classified into 2 categories:

a) a single abnormal result
b) persistent antibodies (2 independent antibody samples conducted ≥2 times more than 14 days apart)

No assessment of neutralising antibodies was undertaken for alemtuzumab. The presence of neutralising antibodies in other phases could be entered as freetext if this was a reason for treatment discontinuation.

JC virus status / index

JC virus serology is a major determinant of PML risk and was documented when available during natalizumab treatment, including any titre or index result where available.

Statistical Analysis

Demographic and background information was summarized using frequency distributions (for categorical variables) and descriptive statistics of mean, standard deviation, minimum, median and maximum (for continuous variables) where appropriate. Relevant medical history/current medical conditions were summarised. Concomitant medication and non-drug therapy were summarised and categorised in each phase.

Duration (months) of exposure to alternative disease modifying therapy (during phase 1, 3 or 5), natalizumab (during phase 2) and alemtuzumab (during phase 4) was also tabulated.

Primary analysis

Descriptive statistics were used to report the primary and secondary endpoints described above.

Adverse events and serious adverse events were summarised by presenting the number and percentage of patients having any adverse event, and grouped as defined above.

Mean annualised relapse rate, number and proportion of patients with sustained accumulation (SAD) or reduction (SRD) of disability, and number and proportion of patients developing significant MRI markers of active inflammatory disease during each of the 5 phases are presented. The number and proportion of patients free of relapse, significant MRI markers of active disease, or sustained accumulation of disability ('freedom from disease activity') was also evaluated in each phase. ARR was calculated as the total number of relapses in a phase divided by the total number of patient-years in that phase, rather than using each individual patient's ARR, given the variable duration of each phase for each patient and the skewing to higher rates which would have occurred in those with shorter duration phases.

Disability outcomes (EDSS scores) were also assessed using the area under the curve method[236] during each treatment phase. The area under the EDSS/follow-up time curve (AUC) was calculated as per the trapezium method correcting for baseline disability and rescaling—with changes of 0.5 for EDSS scores ≥5.5 and <7.0 being normalised to a 1.0 point change at all other levels of the scale. Individual patients were then categorised as (i) 'net improved' for an AUC of <−0.5 EDSS-years; (ii) 'net worse' for an AUC > +0.5 EDSS-years and (iii) 'net unchanged' for an AUC between -0.5 and 0.5 EDSS-years[237]. Patients with available data were also categorised into one of six descriptive disease course categories from the profile of their plot of EDSS-change versus follow-up time. Under this classification system, SRD and SAD become 'sustained improvement' and 'sustained progression' if the disability change is maintained until last follow-up or, if not, as 'erroneous improvement' and 'erroneous progression'[36]. 'Minimal change' indicates an EDSS change ≤0.5 points from baseline at all measurements made over the course of follow-up. The remaining profiles that do not fit any of the above categories are labelled as 'fluctuating'. From these six categories, three groups were defined: 'confirmed stable' ('sustained improvement' or 'minimal change'); 'unsustained change' ('erroneous progression', 'erroneous improvement' or a 'fluctuating' course); and

'confirmed worsening' comprising those with 'sustained progression'. The time point (treatment phase) at which disability changes occur in each patient leading to their classification in this manner is assessed.

Secondary progression is defined as two consecutive SAD events, the second from the new EDSS baseline established after the first SAD event, and in which the increase in EDSS occurred independent of relapses.

Secondary analysis

Comparable descriptive statistics will be presented for all secondary variables.

Ethical requirements

The protocol proposal was assessed by the West of Scotland Research Ethics Service (based in the Tennant Institute in Glasgow) as a phase IV study, not requiring ethical approval. Discussions with, and approval from, the Caldicott Guardian or Clinical Audit Lead (in Ireland) were advised in each centre, initiated by the centre lead. Patients were identified by a unique identifier and no patient identifiable data was held on the online database. All clinical and radiographic data was handled exclusively by the local clinicians (who will ordinarily have been the patient's treating team) and the blinded neuro-radiologist in each centre. The study posed no clinical risk to patients, and did not infringe on their rights or confidentiality as per the approval of the Glasgow Caldicott Guardian, for example. No additional data was collected on patients over and above the clinical data already in local files.

Funding

The study was funded by Sanofi-Genzyme® pharmaceuticals. NHS Greater Glasgow & Clyde was the Sponsor and a contract was agreed between both parties. Sub-sites were permitted for inclusion by way of agreement to a Letter of Comfort by the centre lead. Importantly, data access by Sanofi-Genzyme® and TechNeo® is limited to the overall results rather than patient-level data, as indicated by the Data Flow Map in **Appendix 5**.

Results

Cohort Characteristics

A total of 79 patients with RRMS who had switched from natalizumab to alemtuzumab were identified from 13 MS centres across the UK and Ireland. The number of patients per location is outlined in **Table 3-1**.

Table 3-1: Location of patient centres

MS Centre	Number of patients
Glasgow	5
Newcastle	1
Sunderland	5
Dublin	14
Leeds	5
Liverpool	1
Sheffield	1
Birmingham	12
Swansea	1
Cardiff	1
Bristol	5
Charing Cross	19
Plymouth	9
Total	**79**

For safety analysis, the entire cohort was used (N=79) and for efficacy analysis the cohort with follow-up for at least 2 years after the first course of alemtuzumab was used (N=51). Similarly, MRI data were available for 36 patients (MRI Cohort) and 25 patients had both MRI and clinical data (MRI Efficacy Cohort). **Table 3-2** outlines the characteristics of each cohort relevant to each study phase.

Given the reliance on medical notes, often from some years previously, data were not available for all variables, as reflected in **Table 3-2**: the number of patients with available data (n) is included where this was not the entire cohort. For example, for CSF and Pre-natalizumab MRI, only 22 patients had available data. There were no statistically significant differences between the Baseline characteristics of the cohorts, using ANOVA for parametric comparisons, Kruskal-

Wallis for non-parametric comparisons and chi-square test for comparisons of proportions between the 4 cohorts.

The final date for duration of follow-up was the end of data collection for the study: 31/12/17 for clinical data and 31/1/18 for MRI data. Datapoints (and their date of collection) were only entered where an event occurred e.g. relapse or EDSS, therefore the (final) date of data entry would underestimate follow-up if no further events occurred hence the final date of study data collection was used. Using the final date of data being entered as the end of follow-up reduced the duration of follow-up for the entire cohort (N=79) to median 1.6 years (range -2.6 to 7.6), with 19 patients having last data entry before alemtuzumab started. In the efficacy group (N=51), using the last date of data entry as the end of follow-up, reduced this to median 2.9 years (range 0 to 7.6) with 2 patients having no data entered after alemtuzumab treatment. The design of the database therefore makes the assumption that patients were followed-up until the end of study data collection and that the absence of data entry reflects an absence of available data or clinical events.

The cohorts were predominantly young and female with onset of their first demyelinating symptoms within months before diagnosis (**Table 3-2**). Despite this, EDSS was greater than 2 on average at diagnosis, with most having had 2 relapses prior to diagnosis. Paraclinical results were in keeping with MS, the majority having few white cells in CSF with normal protein, unpaired oligoclonal bands and most had abnormal VEPs - a small proportion having bilateral delay in optic nerve response. In keeping with the working age of the cohort, the vast majority were employed and driving at the time of diagnosis.

Prior to commencing natalizumab, the majority had been treated with one or two DMTs, most ultimately failing treatments due to lack of efficacy. That said, for over a third of patients natalizumab was their first DMT, suggesting the high disease activity of this cohort given its use as a first-line agent only in patients with rapidly evolving severe MS. The annualised relapse rate before starting natalizumab was relatively low at approximately 0.5 but, by the time the of starting (approximately 4 years after diagnosis on average), average EDSS was notably elevated at 3.8 in this young cohort. Additionally, the majority had developed a moderate or marked supratentorial lesion load in the opinion of our

blinded neuroradiologists retrospectively reviewing imaging undertaken before natalizumab was started. As expected at this early disease stage, there was not significant evidence of brain atrophy on visual inspection of MRI imaging however.

Patients started natalizumab in their early 30s on average, approximately 4.5 years after their first demyelinating event had occurred. Natalizumab was used for 1-2 years for most patients, a median of 14 infusions in total. Unusually, given its evidence-based place as a highly effective DMT, the majority of patients failed natalizumab due to efficacy but concerns about PML were also clearly prevalent and a high proportion were JCV antibody positive. A small proportion had neutralising antibodies to natalizumab found but these were not routinely measured in all patients.

The switch period between natalizumab and alemtuzumab lasted 8 months overall and almost 6 months for the efficacy cohort on average. The vast majority of patients did not use DMTs during the switch period. This was not centre-specific, with patients using DMTs during the switch period being from one of seven centres in the total cohort and from four different centres in the efficacy cohort.

The efficacy cohort by definition had a longer duration of follow-up data available after alemtuzumab was started (mean 4.8 years) in comparison to the cohort overall (mean 3.6 years) but the longest follow-up was available for the MRI efficacy cohort (mean 5.3 years). The majority of patients had 2 courses of alemtuzumab in total, as expected from its dosing guidance. Now in their mid-30s on average by the time of alemtuzumab initiation, patients had a disease duration of around 7 years at the time of starting this treatment.

The vast majority did not require further DMTs after using alemtuzumab but, in those who did, further courses of alemtuzumab were used in most. In the efficacy cohort with available follow-up data, 16 patients (31%) had DMTs after using Alemtuzumab, most using one only thereafter (n=12, 75%) but 1 patient requiring 2 further DMTs and 3 patients (19%) using 3 DMTs after alemtuzumab. Further alemtuzumab treatment due to ongoing disease activity after our pre-defined 2-year treatment phase was required by 14 patients (27.4%). Three patients (5.8%) had a further course of alemtuzumab, having had only 1 course

of treatment initially, 10 (19.6%) had a 3rd course and 1 patient (1.9%) required 4 courses in total. Additionally, fingolimod was used in one patient due to ongoing disease activity despite 3 courses of alemtuzumab.

Haematopoietic Stem Cell Transplantation was undertaken for 3 patients with ongoing disease activity despite treatment with natalizumab and alemtuzumab. Of these, 2 were from a single centre. Disease activity continued despite a 3rd course of alemtuzumab in 2 of these patients and one was treated with cyclophosphamide for 3 months before HSCT was undertaken. Of these 2 patients, one actually went on to have a second HSCT 14 months after the first and the second patient was treated with rituximab after HSCT but this was for a post-transplant lymphoproliferative disorder.

Table 3-2: Cohort characteristics

	Total (Safety) Cohort (N=79)	Efficacy Cohort (N=51)	Total MRI Cohort (N=36)	MRI Efficacy cohort (N=25)
Baseline characteristics				
% Female	55 (69.6%)	40 (78%)	30 (83%)	21 (84%)
Mean Age at diagnosis (SD)	31 (9)	29 (8)	29 (8)	28 (7.3)
Median Disease duration at diagnosis [Years] (IQR)	0.37 (1.1)	0.34 (1.1)	0.11 (0.61)	0.07 (0.38)
Mean baseline EDSS (SD)	2.4 (1.7) [n=59]	2.3 (1.7)	2.9 (1.8)	2.8 (1.9)
Mean number of relapses pre-diagnosis (SD)	1.8 (1.3) [n=61]	1.9 (1.3)	1.7 (1.0)	1.8 (1.0)
CSF WCC <5 5-20 >20 (max 40)	[n=22] 13 (59%) 8 (36%) 1 (4.5%)	[n=20] 12 (60%) 7 (35%) 1 (5%)	[n=7] 4 (57%) 2 (29%) 1 (14%)	[n=6] 3 (50%) 2 (33%) 1 (17%)
CSF Protein ≤0.5 >0.5 (max 0.77)	[n=22] 12 (55%) 10 (45%)	[n=20] 11 (55%) 9 (45%)	[n=9] 5 (56%) 4 (44%)	[n=8] 4 (50%) 4 (50%)
CSF OCBs Normal Unpaired Paired	[n=28] 1 (4%) 23 (82%) 4 (14%)	[n=21] 0 18 (86%) 3 (14%)	[n=9] 0 9 (100%) 0	[n=7] 0 7 (100%) 0
VEPs Normal Unilateral delay Bilateral delay	[n=16] 5 (31%) 8 (50%) 3 (19%)	[n=14] 5 (36%) 7 (50%) 2 (14%)	[n=5] 2 (40%) 3 (60%) 0	[n=5] 2 (40%) 3 (60%) 0
Patients with no co-morbidities	63 (79.7%)	39 (76.5%)	27 (75%)	19 (76%)
Employed	38 (76%) (n=50)	27 (66%) (n=41)	17 (59%) (n=29)	13 (72%) (n=18)
Driving	32 (78%) (n=41)	18 (78%) (n=23)	13 (68%) (n=19)	6 (75%) (n=8)
Pre-Natalizumab				
Number of DMTs used (%) 0 1 2 3 4 5	 34 (43%) 29 (37%) 10 (13%) 4 (5%) 1 (1%) 1 (1%)	 20 (39%) 22 (43%) 4 (8%) 3 (6%) 1 (2%) 1 (2%)	 15 (42%) 15 (42%) 4 (11%) 1 (3%) 0 1 (3%)	 9 (36%) 11 (44%) 3 (12%) 1 (4%) 0 1 (4%)

Main reason for stopping DMT(s) (%)	Efficacy (52%) Side effects (38%)	Efficacy (49%) Side effects (40%)	Efficacy (50%) Side effects (40%)	Side effects (46%) Efficacy (42%)
Mean duration of Pre-Natalizumab phase [years] (SD)	4.3 (4.3)	4.2 (4.4)	3.7 (3.3)	3.7 (3.4)
Mean ARR¶	0.53	0.57	0.55	0.45
Mean EDSS before Natalizumab treatment (last available) (SD)	3.8 (2.0)	3.4 (2.0)	3.9 (2.2) [n=27]	3.6 (2.2)
MRI Brain	(n=22)	(n=15)	(n=22)	(n=20)
Mean Gd+ lesions (SD)	4.9 (10.6)	7.3 (13.5)	4.9 (10.6)	6.9 (12.4)
<u>Supratentorial lesion load (%)</u>				
None	0	0	0	0
Minor	2 (9%)	2 (13%)	2 (9%)	3 (15%)
Moderate	2 (9%)	2 (13%)	2 (9%)	2 (10%)
Marked	18 (82%)	11 (73%)	18 (82%)	15 (75%)
<u>Infratentorial (%)</u>				
None	4 (18%)	4 (27%)	4 (18%)	4 (20%)
Minor	7 (32%)	2 (13%)	7 (32%)	4 (20%)
Moderate	7 (32%)	5 (33%)	7 (32%)	5 (25%)
Marked	4 (18%)	4 (27%)	4 (18%)	7 (35%)
<u>Global atrophy (%)</u>				
None	10 (45%)	9 (60%)	10 (45%)	13 (65%)
Minor	10 (45%)	6 (40%)	10 (45%)	3 (15%)
Moderate	2 (9%)	0	2 (9%)	0
Marked	0	0	0	0
<u>Callosal atrophy (%)</u>				
None	11 (50%)	8 (53%)	11 (50%)	12 (60%)
Minor	6 (27%)	4 (27%)	6 (27%)	5 (25%)
Moderate	3 (14%)	2 (13%)	3 (14%)	2 (10%)
Marked	2 (9%)	1 (0.7%)	2 (9%)	1 (5%)
Natalizumab Treatment				
Mean age at 1st Natalizumab Treatment (SD)	33 (9.4)	34 (8.9)	33 (8.6)	32 (8.5)
Mean disease duration at 1st Natalizumab Treatment* [years] (SD)	4.5 (4)	4.6 (4.4)	4.2 (3.2)	4.2 (3.5)
Median duration Natalizumab Treatment [years] (IQR)	1.8 (2.7)	1.2 (2.1)	0.9 (1.5) [n=19]	0.9 (1.5) [n=19]

Median number of Natalizumab infusions (IQR, range)	14 (33, 2-104) (26 missing)	12 (19, 2-91)	12 (33, 2-91) (n=25)	12 (19, 2-91) (n=19)
Main reasons for stopping Natalizumab	Efficacy (38%) PML concern (16%)	Efficacy (58%) Adverse events (19%) PML Concern (19%)	Efficacy (56%) PML concern (28%) (n=25)	Efficacy (56%) PML (24%)
Number of JCV +ve patients	23 (72%) (n=32)	16 (70%) (n=23)	10 (71%) (n=14)	10 (71%) (n=14)
Number of patients with Natalizumab Neutralising Abs	7 (12%) (n=59)	6 (14%) (n=44)	4 (15%) (n=26)	4 (17%) (n=24)
Switch period				
Mean duration Switch period [days] (SD)	241 (231) [8 months]	174 (150) [5.8 months]	238 (219) [7.9 months]	157 (1135) [5.2 months]
Patients using DMTs during Switch (%)	18 (23%)	7 (14%)	9 (25%)	4 (16%)
Alemtuzumab Treatment				
Mean duration of follow-up after 1st Alemtuzumab dose (years) (SD)	3.6 (2.4)	4.8 (2.1)	4.2 (2.5)	5.3 (2.1)
Mean number of Alemtuzumab courses (SD, range)	2 (1, 1-4) (n=59)	2 (1, 1-4) (n=45)	2 (1, 1-4) (n=25)	2.4 (0.6,2-4) (n=24)
Number of Alemtuzumab Courses 1 2 3 4 Unknown	 13 (17%) 33 (42%) 12 (15%) 1 (1%) 20 (25%)	 2 (4%) 30 (59%) 12 (24%) 1 (2%) 6 (12%)	 1 (3%) 16 (44%) 7 (19%) 1 (3%) 11 (31%)	 0 16 (64%) 7 (28%) 1 (4%) 1 (4%)
Mean age at first Alemtuzumab dose (SD)	35 (9.6)	36 (8.6)	35 (8.7)	33 (8.3)
Mean disease duration at 1st Alemtuzumab dose* (years) (SD)	7.2 (4.5)	6.7 (5.2)	7.5 (4.3)	6.7 (4.8)
Number of patients using DMTs after Alemtuzumab	16 (32%) (n=50)	16 (31%)	9 (34%) [n=26]	9 (36%)

¶ARR calculated by total number of relapses in phase / patient-years in phase

*Disease duration = since first symptom (not diagnosis)

The nature of the first reported MS-related symptom for the cohorts are outlined in **Table 3-3**. The predominantly sensory relapses or optic neuritis are in keeping with typical RRMS populations.

Table 3-3: First Demyelinating event symptoms

First demyelinating event	N=79 (Total cohort)	N=51 (Efficacy Cohort)
Pure Sensory	18	15
Optic neuritis	11	9
Pure motor	7	3
Cranial nerve	4	4
Brainstem	4	1
Spinal	4	3
Other visual	3	0
Cerebellar	3	3
Polysymptomatology	3	2
Sensorimotor	2	1
Bulbar	1	1
Unknown	9	9

Whilst the cohorts described are historical, the majority of patients in the total cohort and almost half of the efficacy cohort were treated with alemtuzumab during or after 2015, making them relatively contemporary to current populations (See **Figures 3-2** and **3-3**). Most patients were diagnosed with RRMS in the mid-2000s however, before the updated McDonald 2010 diagnostic criteria (n=57, 72%), and 23 patients (29%) were diagnosed before the 2005 criteria were available. Natalizumab was largely used after its UK licensing date in 2006 but, as anticipated, a large proportion of patients in this cohort (n=33, 42%) were treated with alemtuzumab before its UK licensing date in late 2014, making it unique in the worldwide experience.

Figure 3-2: Year of Diagnosis, Natalizumab and Alemtuzumab initiation in the whole cohort

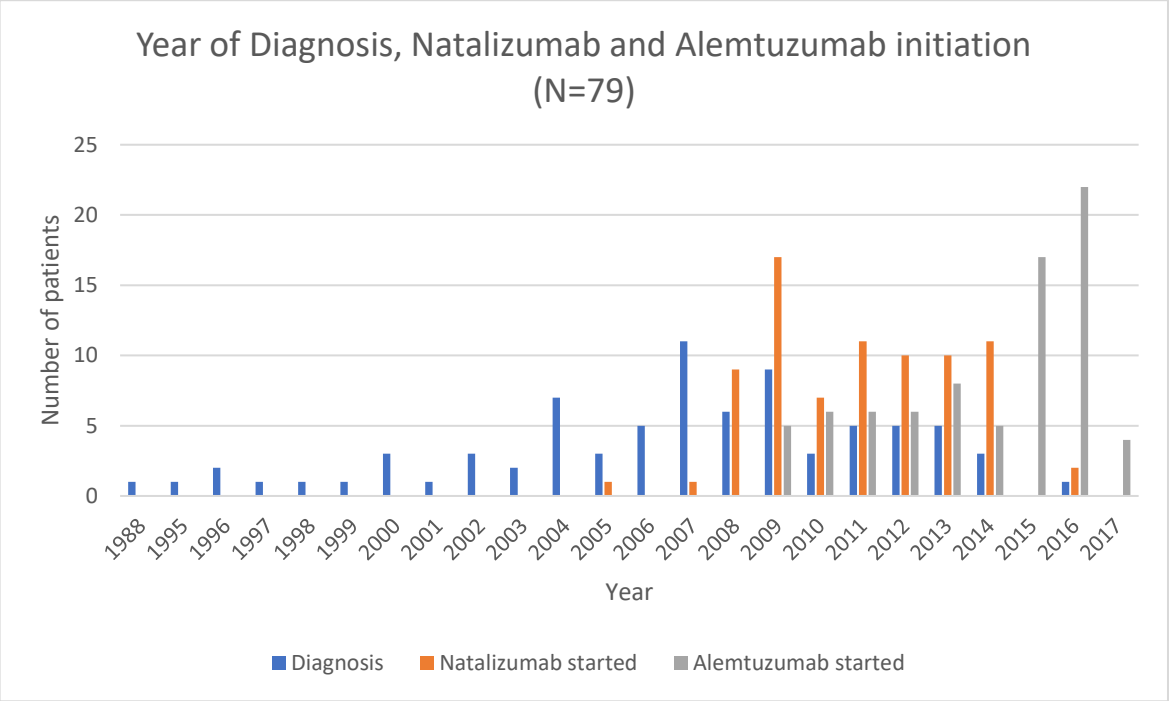
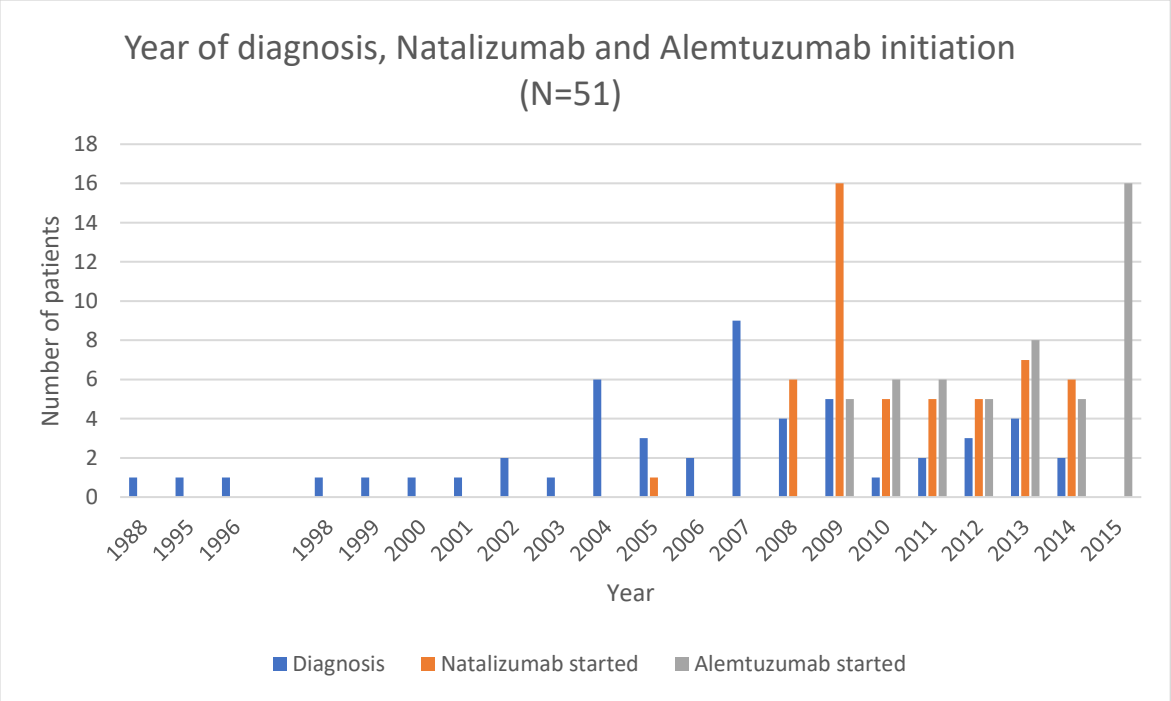


Figure 3-3: Year of Diagnosis, Natalizumab and Alemtuzumab initiation in the efficacy cohort



Co-morbidities

At the time of diagnosis of RRMS, the majority of patients had no other co-morbidities. This was the case for 63/79 (80%) of the total cohort and 39/51

(77%) of the efficacy cohort. In the whole cohort, 16 (20%) had one, 6 (8%) had two and 3 (4%) had three. In the efficacy cohort, 12 (24%) had one, 5 (10%) had two and 3 (6%) had three co-morbidities. The co-morbidities are detailed in **Table 3-4**.

Table 3-4: Co-morbidities by phase

Co-morbidity	Baseline	Pre-Natalizumab	Natalizumab Treatment	Switch Period	Alemtuzumab Treatment	Post-Alemtuzumab
Asthma	3					
Hypertension	2					
Hypothyroidism	2			2		
Bipolar disorder	1		1			
Depression		1			1	
Cholecystectomy	1					
Congenital aortic stenosis	1					
Epilepsy	1				1	
IDDM	1					
Irritable Bowel Syndrome		1				
Joint hypermobility syndrome	1					
Myocardial infarction	1					
Migraine	1					
Parotid swelling	1					
PCOS	1					
Raynaud's phenomenon	1					
Appendicitis	1		1			
Eczema	1					
Glandular fever	1					
Osteoarthritis	1					

Obesity	1	1	1			
Herpes Zoster	1	1				
Allergic rhinitis	1					
Labial cyst		1				
Retinal detachment		1				
Urticaria		1	2			1
Infection NOS		1	1			1
Eosinophilia			1			
Erectile dysfunction			1			
Pilondal sinus			1			
UTI			2			2
Constipation			1			
Abdominal pain NOS			1			
Neurogenic bowel			1			
Osteoporosis			1			
Unilateral deafness			1			
Shoulder pain			1			
Lichen Simplex			1			
Intertrigo			1			
CMV infection					1	
Gastritis					1	
Rash NOS					1	
Lymphoedema						1
Post-Tx lymphoma						1
Hyperthyroidism						1
Tonsillitis						1
Norovirus						1

None	63	69	64	74	70	68

After baseline there were missing co-morbidities data, as outlined in **Table 3-5**. The large amount of missing data makes comparisons in each phase difficult. Only 13 patients had co-morbidities listed for all phases and 11 of these had no co-morbidities throughout. With this caveat, the mean number of co-morbidities per phase is also included in **Table 3-5** and appear to increase over time as would be expected with age, but the influence of MS and its treatment cannot be clarified in the absence of a control group.

Table 3-5: Number of co-morbidities by phase in the efficacy cohort

Phase	No of patients with data	No. Missing (%)	Mean number of Co-morbidities (±SD)
Baseline	51	0	0.41±0.85
Pre Natalizumab	23	28 (55%)	0.65±0.93
Natalizumab	22	29 (58%)	0.91±1.06
Switch	16	35 (69%)	0.50±0.89
Alemtuzumab	19	32 (63%)	0.58±0.84
Post Alemtuzumab	21	30 (59%)	0.86±1.24

Disease modifying therapies

The majority of patients used DMTs prior to natalizumab but only a minority used DMTs during the Switch period or Post-alemtuzumab. **Table 3-6** shows the use of DMTs in these phases for the efficacy cohort. The increase in the proportion of patients with no DMT from Pre-natalizumab (39%) to Post-alemtuzumab (69%) was significant ($p<0.01$). This suggests that need for DMTs was significantly reduced after patients were treated with natalizumab and alemtuzumab.

The increase in annualised DMT rate (per 100 patient years) between Pre-natalizumab (22.8) and Switch (35.1) and the fall to Post-alemtuzumab (16.8) were not significant.

Table 3-6: Disease modifying treatments (DMT) by phase

Phase (N=51)	Pre-Natalizumab	Switch	Post-Alemtuzumab
No of patients with no DMT (%) (95% CI††)	20/51 (39%) (27%,53%)	44/51 (86%) (74%,93%)	35/51 (69%) (55%,80%)¶
Total number of DMT (max per patient)	49 (5)	8 (2)	24 (3)
Total patient yrs in phase	215.2	22.8	142.9
Overall annualised DMT rate* (95% CI†)	22.8 (16.8,30.1)	35.1 (15.1,69.1)	16.8 (10.8,25.0)

*Per 100 patient years †Using Byar's method ††Using the Wilson Score method

¶p<0.01 vs Pre-Natalizumab phase

Pre-Natalizumab phase

As outlined in **Table 3-2**, the majority of patients were treated with DMTs before starting natalizumab (n=45, 57%). **Figures 3-4** and **3-5** outline the number and type of DMTs used during this period respectively. Of those treated with a DMT, most had one or two, but up to 5 were used in one patient. The injectable therapies were used for almost 90% of patients in the cohort and typically sequencing between these, rather than escalating to a higher efficacy treatment, was undertaken prior to use of natalizumab.

Figure 3-4: Number of DMTs used Pre-natalizumab

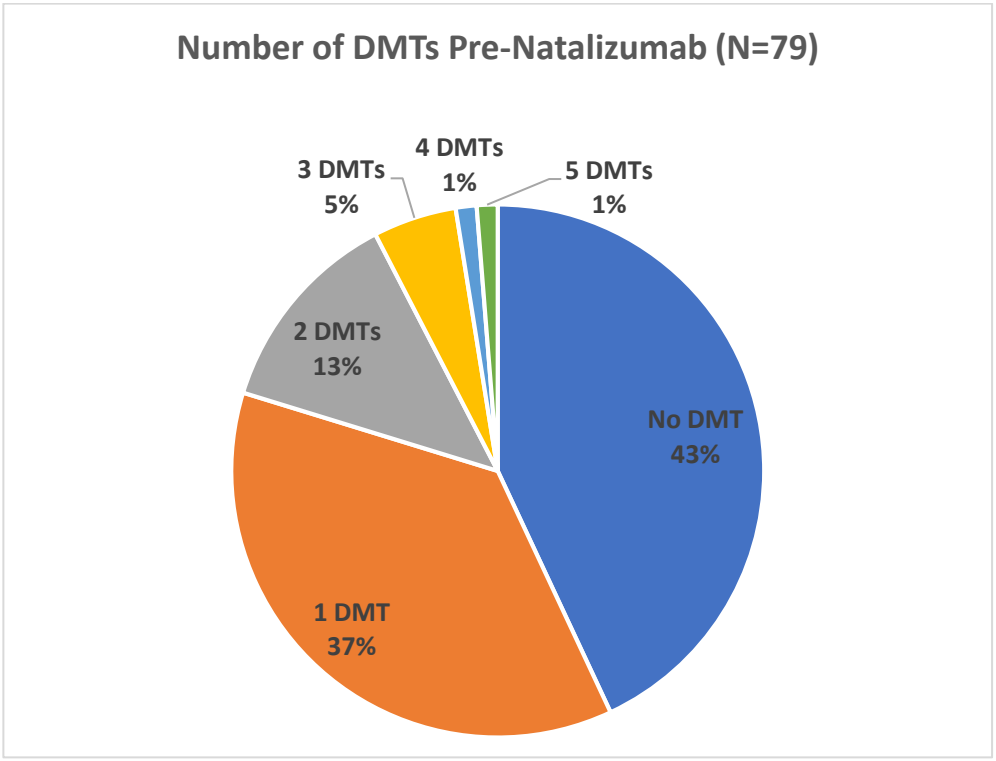
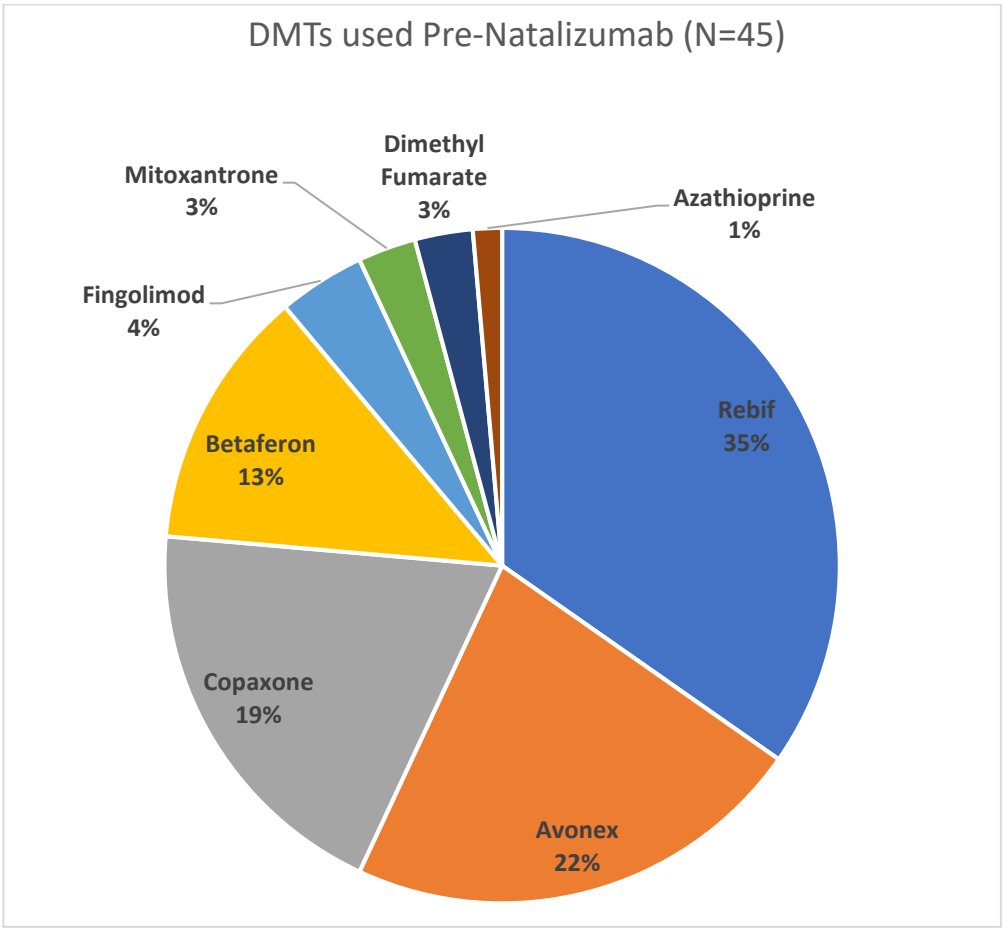


Figure 3-5: DMTs used Pre-natalizumab



Natalizumab Treatment Phase

Most patients switched from natalizumab to alemtuzumab because of treatment failure, either clinical, radiological or both. Use of natalizumab in the UK is restricted to patients having 2 disabling relapses in the past year and with significant increase in MRI disease activity, known as Rapidly Evolving Severe MS (RES MS), meaning these patients had active disease in order to receive this treatment. One case was documented as switching to alemtuzumab solely for increased radiological disease activity but further detail beyond treatment failure due to disease activity was not available for all cases. A notable proportion of patients did not have a reason recorded in the study/medical notes, however, as outlined in **Table 3-7** and **Figure 3-6**. Adverse events were predominantly recorded as hypersensitivity reactions, in one case a SAE requiring hospital treatment. Notably, this patient had positive natalizumab neutralising antibodies but the others with hypersensitivity events did not. Other adverse events included one patient with recurrent infections during treatment and one with persistent cough. The third patient discontinuing due to an adverse event did not have available details on the nature of this.

Concerns about the development of PML, presumably both patient and clinician, also accounted for a significant proportion of patients switching from natalizumab to alemtuzumab. **Table 3-8** outlines the JCV antibody results for the entire cohort, measured only during the natalizumab phase. As expected, all patients who switched to alemtuzumab from natalizumab due to concerns about PML were JCV positive. JCV titres were documented for just 3 patients positive for the antibody: these were 0.44, 0.93 and 4.43 respectively.

Table 3-7: Reason for switch from Natalizumab to Alemtuzumab

	n (N=79) Total Cohort	n (N=51) Efficacy Cohort
Efficacy	30	25
JC +ve/PML concern	13	8
Hypersensitivity	6	6
Adverse event	3	2
Patient choice*	5	2
Non-compliance	1	1
Neutralising Abs	1	0
Not documented	20	7

*2 found travelling difficult, others not specified

Figure 3-6: Reasons for stopping Natalizumab (N=51)

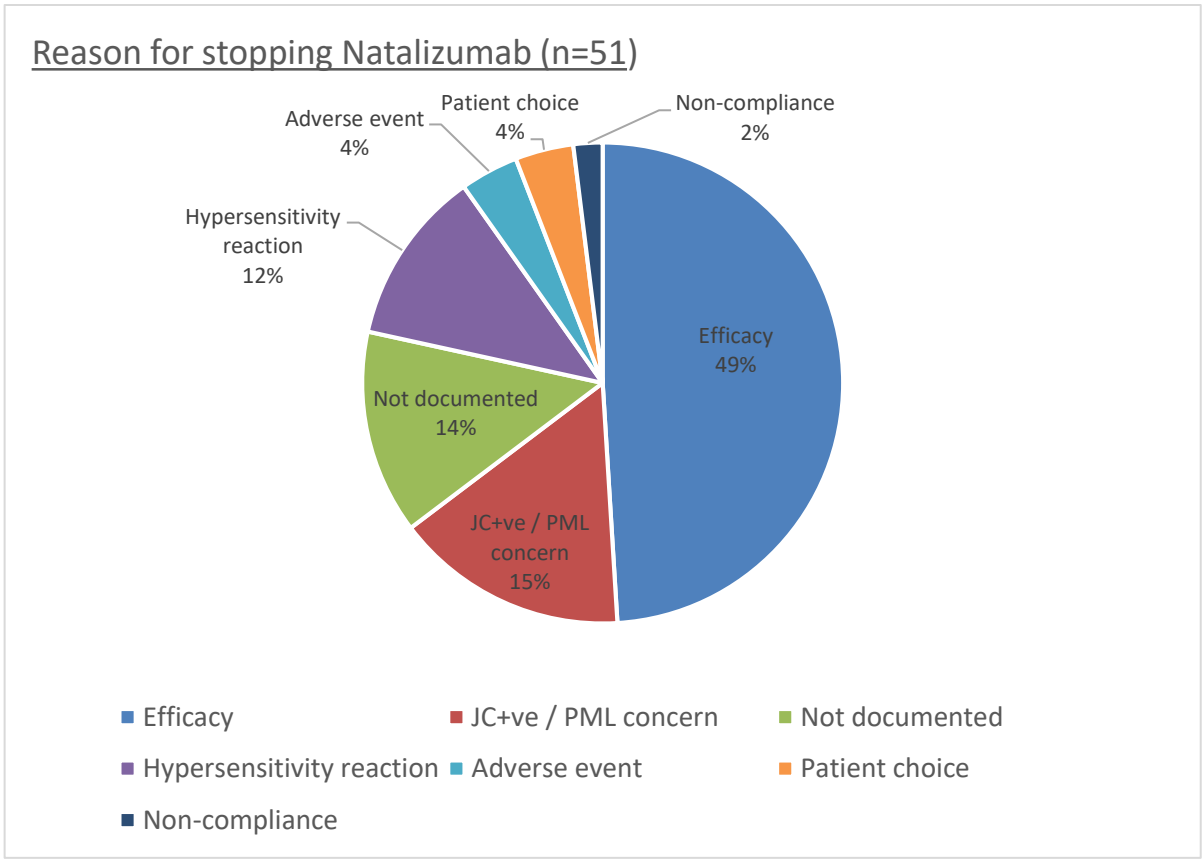


Table 3-8: JCV antibody results during Natalizumab treatment

	N=79
Positive	23 (29%)
Negative	9 (11%)
Not checked in clinical practice	6 (8%)
Not documented in study	41 (52%)

Safety Results

The following primary endpoints will be addressed in this section:

1) The incidence of infections (including opportunistic infections) during each treatment phase
2) The incidence of secondary autoimmunity during each treatment phase
3) The incidence of infusion reactions, hypersensitivity reactions, new blood or urine monitoring events or neoplastic events during each treatment phase

Safety events related to the use of DMTs and the incidence of adverse events leading to treatment discontinuation in each phase will be outlined here. For the Alemtuzumab treatment and Post-alemtuzumab phases all adverse events and serious adverse events were recorded in order to assess the safety of the sequencing of alemtuzumab after natalizumab, the primary objective of this study. Indeed, the design of the data entry system, whilst allowing AEs to be recorded in all phases, was directed towards capturing adverse events during and after alemtuzumab treatment for this reason. In the other phases, only the safety reasons for treatment discontinuation were recorded. In this section, the whole cohort (N=79) is under evaluation unless otherwise stated. The results are summarised in **Table 3-9**.

Table 3-9: Safety Results Summary (N=79)

	Pre-Natalizumab	Natalizumab Treatment	Switch	Alemtuzumab Treatment	Post-Alemtuzumab	**Total**
	Events leading to treatment discontinuation			All events		
Infections	0	1	0	5	8	**14**
Opportunistic infections	0	0	0	1	0	**1**
Secondary autoimmunity	0	0	0	10	4	**14**
Infusion reactions	0	0	0	48	0	**48**
Hypersensitivity reactions	2	6	0	1	0	**9**
Blood monitoring event				5	0	**5**
Urine monitoring event				0	1	**1**
Neoplastic events	0	0	0	0	1	**1**
Death	0	0	0	0	1	**1**
SAE	1	1	0	2	4	**8**
Total	**3**	**8**	**0**	**72**	**19**	**102**

Pre-Natalizumab Phase

No infections or secondary autoimmunity were recorded during the Pre-natalizumab phase. Efficacy was the main reason for treatment discontinuation in this phase but, in those stopping treatment due to adverse events, injection-site reactions (2, 3%), hypersensitivity reactions (2, 3%) and flu-like symptoms (1, <1%) were reported as leading to treatment discontinuation. One SAE occurred with Rebif® and was a hypersensitivity reaction. The majority of adverse events were not specified in detail however (20, 69%). **Tables 3-10** and **3-11** outline the data obtained from the Pre-natalizumab treatment phase, leading to DMT discontinuation, including adverse events. The injectable treatments were more frequently stopped due to lack of efficacy rather than adverse events but the opposite was true for oral DMTs.

Table 3-10: Reasons for stopping DMTs in the Pre-Natalizumab phase - Summary

	N	Efficacy (%)	Adverse Event (%)	Other (%)
Avonex®	16	8 (50)	6 (38)	2 (13)
Rebif®	25	16 (64)	8 (32)	1 (4)
Betaferon®	9	2 (22)	6 (67)	1 (11)
Copaxone®	14	9 (64)	3 (21)	2 (14)
DMF	2	0	2 (100)	0
Fingolimod	3	1 (33)	2 (66)	0
Azathioprine	1	1 (100)	0	0
Mitoxantrone	2	0	0	2 (100)
Total	**72**	**37 (51)**	**29 (40)**	**8 (11)**

Table 3-11: Reasons for stopping DMTs in the Pre-Natalizumab phase - Details

DMT		1st-line	2nd-line	3rd-line	4th-line	5th-line	**Total**
Avonex®	**Number of patients**						
		12	3	1	0	0	**16**
	Reason for stopping						
	Efficacy	5	2	1	-	-	**8**
	AE NOS	4	1	-	-	-	**5**
	Injection-site reaction	1	-	-	-	-	**1**
	Pregnancy	2	-	-	-	-	**2**
Rebif®	**Number of patients**						
		20	3	1	0	1	**25**
	Reason for stopping						
	Injection-site reaction	1	-	-	-	-	**1**
	Hypersensitivity reaction	1	-	-	-	-	**1**
	AE NOS	4	-	-	-	-	**4**
	Flu-like symptoms	-	1	-	-	-	**1**
	Efficacy	12	2	1	-	1	**16**
	Patient preference	1	-	-	-	-	**1**
	SAE (hypersensitivity)	1	-	-	-	-	**1**
Betaferon®	**Number of patients**						
		5	3	1	0	0	**9**
	Reason for stopping						
	AE NOS	3	3	-	-	-	**6**
	Pregnancy	1	-	-	-	-	**1**
	Efficacy	1	-	1	-	-	**2**
Copaxone®	**Number of patients**						
		7	4	3	0	0	**14**

	Reason for stopping						
	Efficacy	6	3	-	-	-	**9**
	Part of clinical trial	1	-	1	-	-	**2**
	AE NOS	-	1	1	-	-	**2**
	Worsening mobility	-	-	1	-	-	**1**
DMF	**Number of patients**						
		0	0	1	1	0	**2**
	Reason for stopping						
	AE NOS	-	-	1	1	-	**2**
Fingolimod	**Number of patients**						
		1	1	0	0	1	**3**
	Reason for stopping						
	AE NOS	1	-	-	-	-	**1**
	Efficacy	-	1	-	-	-	**1**
	2nd degree heart block	-	-	-	-	1	**1**
Azathioprine	**Number of patients**						
		0	1	0	0	0	**1**
	Reason for stopping						
	Efficacy	-	1	-	-	-	**1**
Mitoxantrone	**Number of patients**						
		0	1	0	1	0	**2**
	Reason for stopping						
	Reached max dose	-	1	-	-	-	**1**
	Bridging therapy	-	-	-	1	-	**1**

Natalizumab Treatment Phase

Similar to previously-used DMTs, natalizumab was discontinued in most cases because of treatment failure (n=30, 38%) but adverse events (n=9, 11%) were also reported as outlined in **Table 3-7** and **Figure 3-6**. A single SAE was reported with natalizumab, listed as a hypersensitivity reaction. Hypersensitivity reactions were the most common adverse event (n=6, 8%) but one patient developed recurrent infections during treatment and one had persistent cough which resolved on treatment discontinuation.

Switch Phase

During the Switch phase only one adverse event, leading to discontinuation of fingolimod, was recorded but details not specified.

Alemtuzumab Treatment and Post-Alemtuzumab phases

Following alemtuzumab treatment, 85 AEs occurred in total and 37 patients (46.8%) had at least one documented. We estimate that over 575 Alemtuzumab infusions are captured in this dataset but this is limited by some missing data with regards to the dose of each course given (See **Table 3-12**). It is assumed, where not available, that the first alemtuzumab course included 5 infusions and subsequent courses 3 infusions per course.

In the entire cohort, 13 patients developed infections (16%) after alemtuzumab treatment, 3 of which were SAEs (4%) and 1 of which was an opportunistic (CMV) infection. Secondary autoimmune disorders (n=14, 17%) were all thyroid-related with 5 patients (6%) developing hypothyroidism, 8 hyperthyroidism (10%) and 1 (1%) not specified. Of the AEs during the Alemtuzumab treatment phase, the majority [48 (56%)] were infusion reactions. Thrombocytopaenia occurred in 2 patients (3%) and proteinuria and (transient) neutropaenia in 1 each (1%) with none requiring specific therapy. Importantly, the majority of this cohort will not have had the intensive monitoring regime in current practice after alemtuzumab so these will be underestimates.

One patient developed malignancy, related to HSCT following alemtuzumab, and one patient died in the Post-alemtuzumab phase. The death was a female aged 32 with rapidly worsening MS from onset. She was significantly disabled even before alemtuzumab, given in 2013, and died 3 years later from urosepsis which the treating team felt did not directly relate to her DMT use but rather the aggressive course of her disease despite treatment. Secondary autoimmunity occurred both during the 2-year Alemtuzumab treatment phase and after but infections more commonly occurred in the latter in this cohort. The details are outlined in **Table 3-13**.

Table 3-12: Number of Alemtuzumab courses and infusions

Total number of Alemtuzumab courses	N (patients)	Number of infusions
1	13	65
2	33	264
3	12	132
4	1	14
Unknown (but at least 1)	20	≥100
Total	**79**	**≥575**

Table 3-13: Adverse Events in the Alemtuzumab and Post-Alemtuzumab phases (N=79)

AE	Details	Course 1	Course 2	Course 3	Post-Alemtuzumab	**Total**
Infusion reaction	Headache	3	1			**4**
	Rash	22	4	2		**28**
	Arm weakness	1				**1**
	Nausea	1				**1**
	Pyrexia	2	1			**3**
	Tingling in legs	1				**1**
	Leg bruising	1				**1**
	Loose stools	1		1		**2**
	Malaise	1				**1**
	SOB	1				**1**
	Wheeze	1				**1**
	Low back pain		1			**1**
	Chest pain/ tightness	1	1			**2**
	Flu-like symptoms		1			**1**
					Total	**48**
Organ-specific autoimmunity	Hypothyroidism	3	1		1	**5**
	Hyperthyroidism	4	1		3	**8**
	Autoimmune thyroid NOS	1				**1**
					Total	**14**
Hypersensitivity reaction	Urticaria	1				**1**
					Total	**1**
Blood/urine disorders	Persistent lymphopaenia	2				**2**
	Thrombocytopaenia	2				**2**
	Neutropaenia (transient)	1				**1**
	Proteinuria				1	**1**
					Total	**6**
Infections	Shingles	3				**3**

	Oral candida				1	**1**
	Fungal skin infection				1	**1**
	UTIs				2	**2**
	Tonsillitis				1	**1**
	Norovirus				1	**1**
					Total	**9**
Other	Lymphoedema				1	**1**
					Total	**1**
SAEs						
Infection	NOS				2 (1 post-HSCT)	**2**
	UTI	1				**1**
	CMV	1				**1**
Malignancy	Post-Tx lymphoproliferative disorder				1 (post-HSCT)	**1**
Death	Death from urosepsis	Alemtuzumab first dose 18/2/13 age 29; death Feb 2016			1 ('Unrelated to Alem Rx')	**1**
					Total	**6**
Total					**85**	

Adverse events in the efficacy cohort (N=51)

There was a total of 69 adverse events across 30 patients. Sixteen (35%) had no adverse events; 16 (35%) had one event, 9 (20%) had two events, 3 (7%) had 4 events, one patient had 5 and one patient had 6 adverse events. There were more adverse events during the first alemtuzumab course (44 events across 30 patients) than the second course (13 events across 8 patients). All patients with an adverse event during the second course also had an adverse event in the first course. The number of adverse events in patients receiving further alemtuzumab in the Post-alemtuzumab phase was similar to the second course (12 events across 7 patients).

The most common type of adverse events were infusion reactions which accounted for 38/57 (67%) of the events in the Alemtuzumab treatment phase. Organ-specific autoimmunity accounted for 12/69 (17%), Infections 10/69 (14%) and Blood/urine disorders 6/69 (9%) accounted for most of the other adverse events. **Tables 3-14** and **3-15** provide a more detailed breakdown.

Table 3-14: Adverse Events in the Alemtuzumab and Post-Alemtuzumab phases in the efficacy cohort – Summary (N=51)

	Alemtuzumab Course 1	Alemtuzumab Course 2	Post-Alemtuzumab	Total
No. of patients with no adverse events	16/46 (35%)	36/44 (82%)	44/51 (86%)	
No. of patients with adverse events	30/46 (65%)	8/44 (18%)	7/51 (14%)	
Total adverse events	44	13	12	69
Categories				
Infusion reaction	29	9	0	38 (55%)
Organ-specific autoimmunity	6	2	4	12 (17%)
Infections	3	2	5	10 (14%)
Blood/urine disorders	5	0	1	6 (9%)
Other	0	0	2	2 (3%)
Hypersensitivity reaction	1	0	0	1 (1%)

Three of the serious adverse events outlined in **Table 3-13** occurred in patients in the efficacy cohort. The patients with urinary tract and CMV infections SAEs and the patient who died were included in this cohort.

Table 3-15: Adverse Events in the Alemtuzumab and Post-Alemtuzumab phases in the efficacy cohort – Details (N=51)

Adverse events	Alemtuzumab Course 1	Alemtuzumab Course 2	Post-Alemtuzumab	Total
Infusion reaction	**29**	**9**	**0**	**38**
Headache	2	1		3
Rash	17	4		21
Arm weakness	1			1
Nausea	1			1
Pyrexia	2	1		3
Tingling in legs	1			1
Leg bruising	1			1
Loose stools	1			1
Malaise	1			1
SOB	1			1
Wheeze	1			1
Low back pain		1		1
Chest tightness		1		1
Flu-like symptoms		1		1
Organ-specific autoimmunity	**6**	**2**	**4**	**12**
Hypothyroidism	2	1	1	4
Hyperthyroidism	3	1	3	7
Autoimmune thyroid NOS	1			1
Hypersensitivity reaction	**1**	**0**	**0**	**1**
Urticaria	1			1
Blood/urine disorders	**5**	**0**	**1**	**6**
Persistent lymphopaenia	2			2
Thrombocytopaenia	2			2
Neutropaenia (transient)	1			1
Proteinuria			1	1
Infections	**3**	**2**	**5**	**10**
Shingles	3			3
Oral candida		1		1
Fungal skin infection		1	1	2
UTIs			2	2
Tonsillitis			1	1
Norovirus			1	1
Other	**0**	**0**	**1**	**1**
Lymphoedema			1	1

Efficacy Results

The efficacy results are based on the cohort with at least 2 years' follow-up after alemtuzumab was started (N=51) and all following results are based on these data unless otherwise specified.

Relapses

The following endpoints are addressed in this section:

Primary endpoint

- Annualised relapse rate during each treatment phase

Secondary endpoint

- The incidence of clinically significant or disabling relapses during each treatment phase

The annualised relapse rate significantly reduced when DMTs were started after diagnosis and a further significant reduction occurred after alemtuzumab treatment, in comparison to all preceding treatment phases, which was sustained in the Post-alemtuzumab phase. The percentage of patients with no relapses increased significantly in the Alemtuzumab and Post-alemtuzumab phases compared to the Pre-natalizumab and Natalizumab treatment phases. The proportion of patients with no relapse was also higher in the Switch phase but this is likely because it was shorter than the other phases.

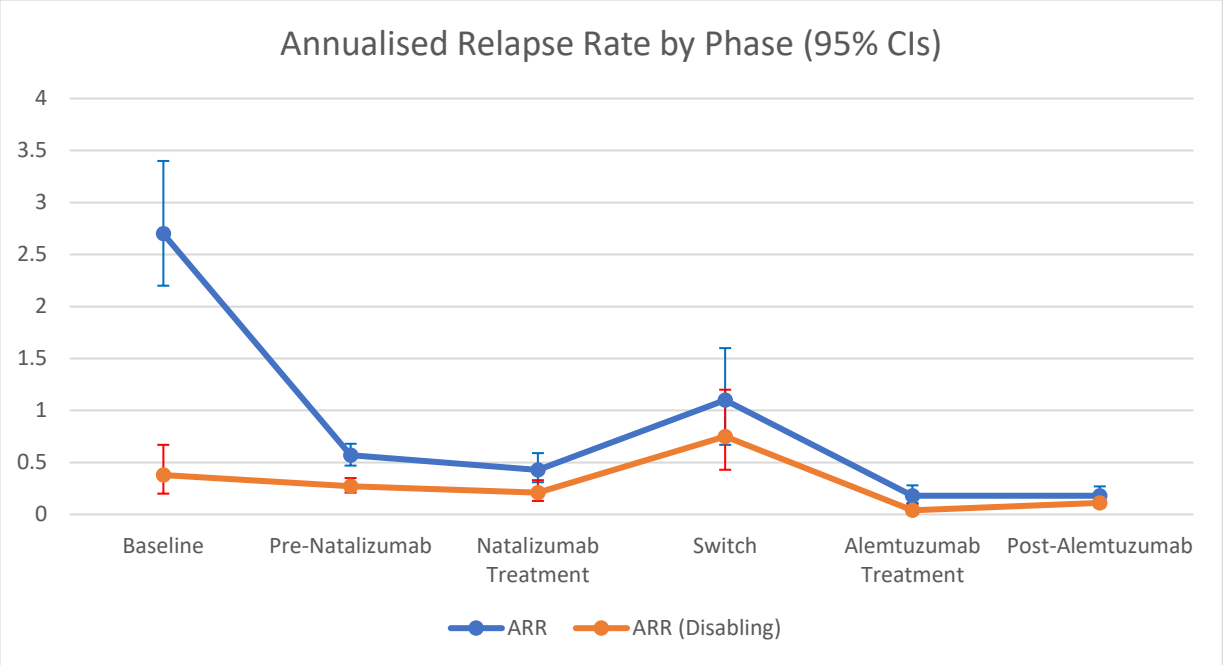
Annualised relapse rates decreased in the Pre-Natalizumab and Natalizumab treatment phases compared to Baseline and increased in the Switch phase before decreasing again in the Alemtuzumab treatment and Post-alemtuzumab phases. The rate in the Alemtuzumab phases was significantly less than in the Natalizumab phase, notably. The same was true for the disabling relapse rate, which was significantly less in the Alemtuzumab phases compared to all other phases, as outlined in **Table 3-16** and **Figure 3-7**.

Table 3-16: Relapses by phase

Phase	Pre-diagnosis (n=44)	Pre-natalizumab (n=51)	Natalizumab treatment (n=51)	Switch (n=51)	Alemtuzumab treatment (n=51)	Post-alemtuzumab (n=50)
No. of patients with no relapses (%) (95%CI††)	0	16 (31%) (20%,45%)	25 (49%) (36%,62%)	36 (71%) (57%,81%)	36 (71%) (57%,81%)	36 (72%) (58%,82%)
Total relapses (max per patient)	85 (8)	122 (14)	41 (4)	24 (4)	18 (2)	26 (6)
Total patient yrs in phase	31.3	215.2	94.5	22.8	101.9	141.7
Overall annualised relapse rate* (95% CI†)	271.6 (217,336)	56.7 (47.1,67.7)	43.4 (31.1,58.9)	105.3 (67.4,157)	17.7 (10.5,27.9)	18.3 (12.0,26.9)
ARR (95% CIs)	2.7 (2.2,3.4)	0.57 (0.47,0.68)	0.43 (0.31,0.59)	1.1 (0.67,1.6)	0.18¶ (0.11,0.28)	0.18¶ (0.12,0.27)
Total disabling relapses (max per patient)	12 (2)	58 (8)	20 (3)	17 (2)	4 (1)	16 (5)
Overall annualised disabling relapse* rate (95% CI†)	38.3 (19.8,67.0)	27.0 (20.5,34.8)	21.2 (12.9,32.7)	74.6 (43.4,119)	3.9 (1.1,10.1)	11.3 (6.5,18.3)
ARR [Disabling Relapses] (95% CIs)	0.38 (0.2,0.67)	0.27 (0.21,0.35)	0.21 (0.13,0.33)	0.75 (0.43,1.2)	0.04¶ (0.01,0.1)	0.11¶ (0.07,0.18)

*Per 100 patient years †Using Byar's method. ††Using the Wilson Score method, ¶$p<0.05$ (vs other phases)

Figure 3-7: Annualised Relapse Rate by Treatment Phase



EDSS SAD & SRD analysis

The following Primary endpoint is addressed in this section:

- The occurrence of sustained accumulation of disability (defined by an increase from baseline of at least one EDSS point or ≥1.5 points if baseline EDSS score was 0) confirmed over 6 months in each treatment phase

Sustained accumulation of disability (SAD) was defined as an increase in EDSS, sustained for at least 6 months, of ≥1.5 EDSS points if the baseline EDSS was 0; ≥1.0 point if the baseline EDSS was ≥1 but <5.5; and >0.5 points if the baseline EDSS was ≥5.5.

Secondary progression was defined as two consecutive SAD events within a phase, the second from the new EDSS baseline established after the first SAD event, and in which the increase in EDSS occurred independent of relapses. In the entire cohort, only one patient showed secondary progression by these criteria in a single phase, the Pre-natalizumab phase. One patient was classified as having SPMS during the Post-alemtuzumab phase by the clinical team.

Sustained reduction in disability (SRD) was defined as a reduction in the EDSS score of ≥1.0 for EDSS scores below 5.5 or 0.5 for baseline EDSS scores above 5.5 sustained for at least 6 months. Analysis for SRD was restricted to those with a baseline EDSS ≥2.0.

The analysis was undertaken using a fixed baseline for subsequent EDSS comparisons i.e. the first EDSS in each phase for phase comparisons and the EDSS at diagnosis for whole study analyses. A re-baselining method was also undertaken, where the baseline comparison EDSS was the preceding EDSS rather than the first in the phase/at diagnosis i.e. re-baselining after each EDSS change. Ultimately the different analyses did not provide significantly different results and the former, fixed baseline, correlated better with the area under the curve EDSS analysis therefore was thought more likely to reflect reality hence is the chosen method for the results presented in this section.

The majority of patients had no SAD or SRD events during each phase. The proportion with no SAD/SRD events was around 70% for all phases except the shorter switch phase where it was higher (See Table 3-17). Where SAD did occur, it was highest in the Pre-natalizumab and Switch phases and lowest in the Alemtuzumab treatment phase, although the difference in annualised SAD rates were not statistically significant between phases. Annualised SRD rate was highest during the Alemtuzumab phase. This was not statistically significant in comparison to preceding phases but the annualised SRD rate did significantly fall in the subsequent Post-alemtuzumab phase, as outlined in **Table 3-17** and illustrated in **Figure 3-8**. These data suggest a trend toward lower SAD and higher SRD rates during the 2-year Alemtuzumab treatment phase but this is not sustained beyond this.

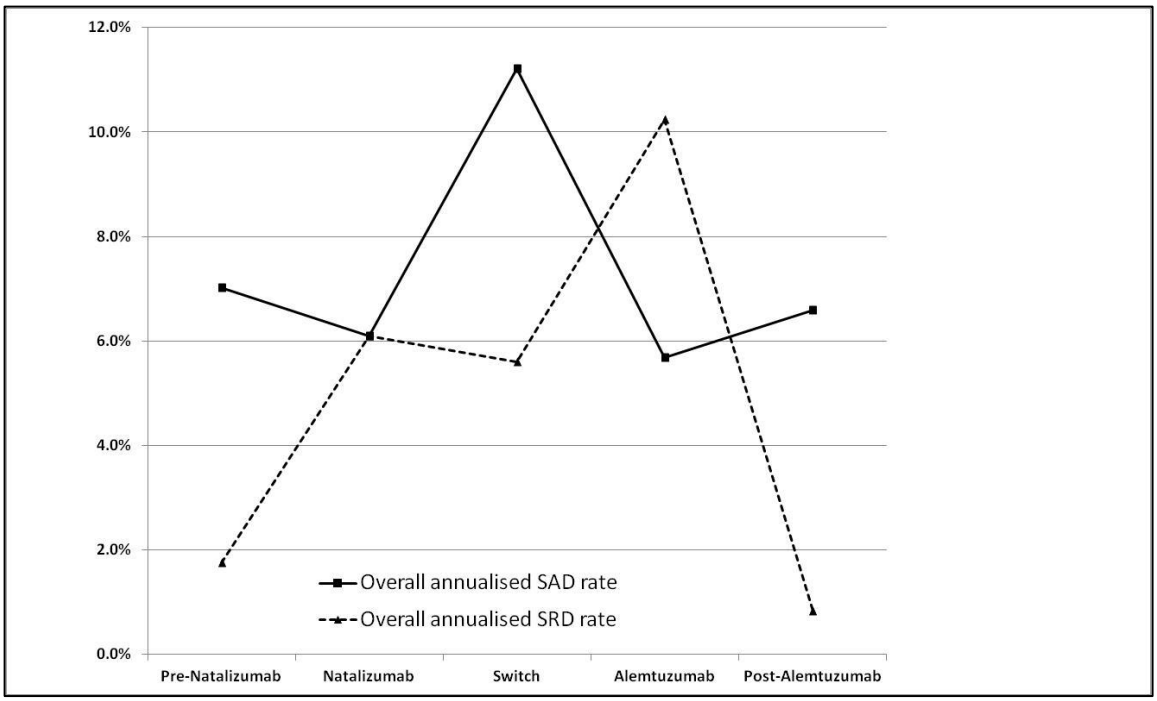
Table 3-17: SAD/SRD by phase

Phase (n=51)	Pre-Natalizumab	Natalizumab	Switch	Alemtuzumab	Post-Alemtuzumab
Patients with no SAD/SRD (%) (95%CI††)	26/40 (65%) (50%,78%)	35/45 (78%) (64%,88%)	34/37 (92%) (79%,97%)	30/44 (68%) (53%,80%)	28/37 (76%) (60%,87%)
Total SAD	12	5	2	5	8
Total SRD	3	5	1	9	1
Total patient yrs in phase	170.79	82.02	17.84	87.93	121.35
Overall annualised SAD rate (95% CI†)	7.0 (3.6,12.3)	6.1 (2.0,14.2)	11.2 (1.4,40.5)	5.7 (1.8,13.3)	6.6 (2.8,13.0)
Overall annualised SRD rate (95%CI†)	1.8 (0.4,5.1)	6.1 (2.0,14.2)	5.6 (0.1,31.2)	10.2 (4.7,19.4)	**0.8**¶ (0.02,4.6)
Patients with missing data	11	6	14	7	14

Per 100 patient years †Using Byar's method. ††Using the Wilson Score method

¶p<0.05 vs Alemtuzumab phase

Fig 3-8: Annualised SAD/SRD rates by phase



SAD/SRD categorisation

There were 35 patients with available EDSS data to permit SAD/SRD analyses across all phases. As a method of describing overall cohort outcomes, rather than relating to each phase, these patients were categorised into one of six descriptive disease course categories from their profile of EDSS-change over the entire study. The mean duration of time in the study for these patients was 11.96 years (SD 4.96).

Under this classification system:

1. 'sustained improvement' if SRD is maintained until last follow-up
2. 'erroneous improvement' if SRD is not maintained until last follow-up
3. 'sustained progression' if SAD is maintained until last follow-up
4. 'erroneous progression' if SAD is not maintained until last follow-up
5. 'Minimal change' if an EDSS change ≤0.5 points from baseline for all measurements
6. 'fluctuating' if the profile does not fit any of the above categories

From these six categories, three groups were defined:

1. ‘confirmed stable’ (‘sustained improvement’ or ‘minimal change’);
2. ‘unsustained change’ (‘erroneous progression’, ‘erroneous improvement’ or ‘fluctuating’ course);
3. ‘confirmed worsening’ comprising those with ‘sustained progression’

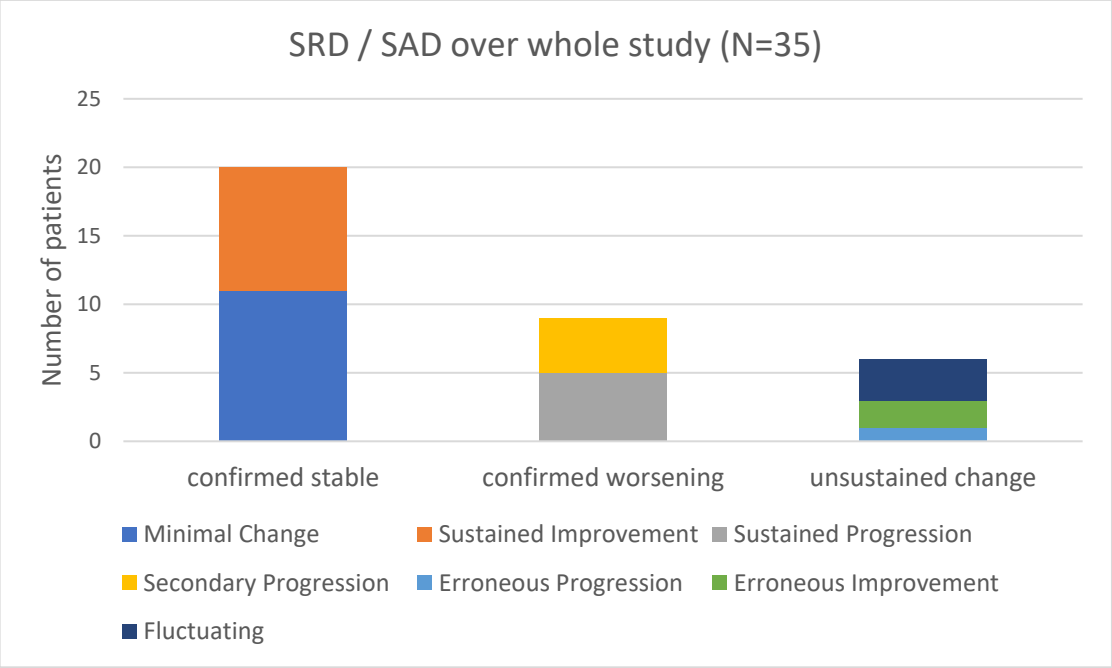
Overall, 20/35 (57%) were confirmed stable, 9/35 (26%) worsened and 9/35 (26%) improved as outlined in **Table 3-18** and **Figure 3-9**. This analysis was dependent on frequently documented EDSS scores in medical notes and the patients included were from one of 8 centres. The lack of EDSS scores in all patients at the required frequency limits conclusions about these results but it would be expected that those with more active disease would have disability scores documented more frequently as a result of increased clinical input. The possibility that the documentation is simply centre-specific is refuted by the number of centres providing these data. With this in mind, the overall SAD/SRD analysis over the whole study period suggests that the majority of patients treated with alemtuzumab after natalizumab remained stable or improved despite being a highly active group with early disability.

Five patients (14.3%) were classified as having secondary progression using this analysis, with two consecutive SAD events over the course of the study irrespective of the treatment phase. Data from natural history studies[12] suggest that by 6-15 years from onset 40% of patients with RRMS will develop SPMS. The much lower figure seen in this cohort suggests a long-term disease-modifying effect from sequenced DMTs, but comparison to historical controls is limited by changing diagnostic guidelines and general progress in healthcare which likely favour more recent cohorts toward improved prognosis. Unknown relevant socioeconomic and/or environmental factors which differ between generations may have either positive of negative effects but also limit such comparisons where these cannot be accounted for.

Table 3-18: SAD/SRD categorisation over entire study (N=35)

SAD/SRD six-fold categorisation		SAD/SRD three-fold categorisation	
	N	N	
Sustained improvement	9	20	Confirmed stable
Minimal change	11		
Erroneous improvement	2	6	Unsustained change
Erroneous progression	1		
Fluctuating	3		
Sustained progression	9	9	Confirmed worsening

Figure 3-9: SRD/SAD categorisation over whole study (N=35)



MRI Results

The following endpoints are addressed in this section:

Primary endpoint

- The occurrence of significant MRI evidence of active inflammatory disease in each treatment phase

Secondary Endpoints

- The occurrence of worsening or improving T2/FLAIR lesions, new T1 lesions and changes in qualitative atrophy during each treatment phase

- To describe any effects of natalizumab and alemtuzumab on qualitative atrophy and cognition in patients with rapidly evolving severe RRMS

There were 36 patients with completed MRI data. The cohort characteristics are outlined in **Table 3-2** but there was significant missing clinical data for a proportion of patients from one centre in particular. Firstly, the total MRI results with all available data will be discussed and then the smaller cohort (N=25) with more complete clinical and MRI data with at least 2 years follow-up after the first alemtuzumab course.

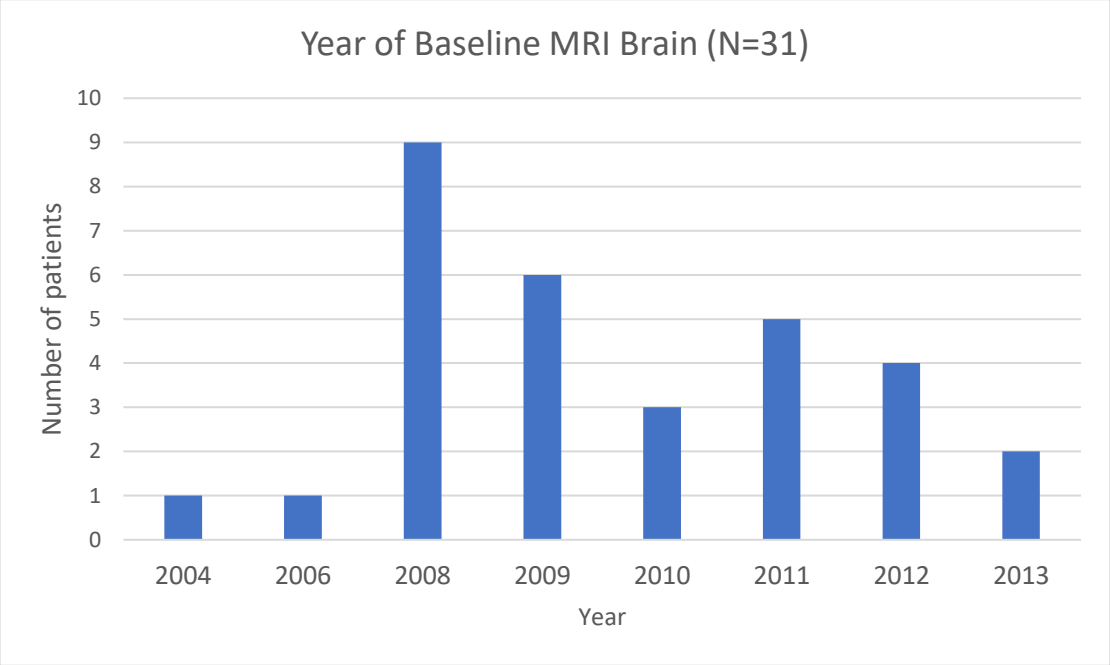
Baseline MRI Brain

31 patients had MRI brain scans from the time of diagnosis or prior to natalizumab treatment as their first available scan, considered their Baseline MRI for this study. The diagnostic scan was available for 9 patients, the other 22 Baseline MRIs were undertaken during the Pre-natalizumab phase, the latter providing the MRI data for **Table 3-2**. **Table 3-19** and **Figure 3-10** detail the centres from which MRI data were available and the year of the Baseline MRI brain scan respectively. Most MRI scans included were from two centres, Birmingham and Dublin, and most Baseline scans undertaken in 2008 and 2009.

Table 3-19: MRI data available by MS Centre

Centre	n
Birmingham	12
Glasgow	4
Dublin	14
Cardiff	1
Leeds	4
Swansea	1
Total	36

Figure 3-10: Year of MRI Brain Baseline (1st available) scan



The reporting blinded neuroradiologist reviewed available scans retrospectively. The number of enhancing lesions were counted on the Baseline scan in those where contrast was given (n=21). Additionally, a qualitative assessment was made on the supratentorial and infratentorial lesion load and degree of both global and callosal atrophy. **Table 3-20** details the findings. There was an average of almost 5 gadolinium-enhancing lesions, but over half of patients had no contrast-enhancing lesions in the 21 scans where this could be assessed. The majority had a marked supratentorial lesion load but minor infratentorial lesion load and minor or no visible atrophy in the opinion of the neuroradiologists.

Table 3-20: Baseline MRI Brain Lesion load and atrophy

Baseline MRI (N=31)				
Mean number of Gd+ lesions* (SD, range)	4.6 (9.7, 0-38)			
Number of patients with no Gd+ lesions* (%)	11 (52%)			
Lesion load qualitative assessment	None	Minor	Moderate	Marked
Supratentorial	0	3	2	25
Infratentorial	4	12	7	7
Atrophy qualitative assessment				
Global	16	12	2	0
Callosal	18	7	3	2

*n=21 (MRIs where contrast was given)

Interval MRI Brain Comparisons

Serial MRI brain scans were compared over time from the Pre-natalizumab phase to the last available scan in the Post-alemtuzumab phase. Only interval scan comparisons which occurred entirely during a phase were included in the final analysis to avoid confounding impacts from treatments outwith that phase. For example, if an MRI undertaken during the Natalizumab treatment phase was compared to a scan undertaken in the Pre-natalizumab phase, based on the dates the scans were undertaken and the beginning and end dates of each phase, this comparison was not included. In a bid to account for this, however, the occurrence of contrast-enhancing lesions was evaluated for each MRI where contrast was given in each phase. The occurrence of contrast enhancement is considered an acute event and therefore negates the concern regarding the impact of previous treatments although we recognise that contrast enhancement can persist for months which may result in false positive events during a treatment phase, although this is relatively rare[238].

New MRI brain lesions occurred most during the Switch and Pre-natalizumab phases, driven mainly by contrast-enhancing lesions in the former and T2/FLAIR lesions in the latter (See **Table 3-21**). The rate of new/worsening lesions was highest during the Switch period despite it being the shortest and with the fewest scans occurring during this time in comparison to other phases (**Figure 3-**

11). The rate of new MRI Brain lesions was lowest in the Natalizumab treatment phase and was also relatively low in the Post-alemtuzumab phase.

The number of improving lesions was highest during the Alemtuzumab treatment phase but the rate per year of phase was highest during the Switch (**Figure 3-12**). Therefore, the Switch phase had the highest rate of both worsening and improving lesions, but the net effect was worsening (**Figure 3-13**). The Switch period also had the highest rate of Gd-enhancing lesions (**Figure 3-14**) and most patients were off DMTs during this phase. From a mechanistic point of view, it could be postulated that stopping natalizumab has resulted in a rebound effect with an open blood-brain barrier permitting both beneficial and deleterious immune activity.

Natalizumab treatment was associated with the lowest rates of new MRI brain lesions and Alemtuzumab with most improved lesions, after the Switch period. Considering both new (worsening) and improving lesions, the net effect favoured the Alemtuzumab treatment period but, again, this was not sustained after 2 years (**Table 3-21**). **Figure 3-13** illustrates the cumulative change in MRI brain lesions throughout each phase, adjusted for the length of each phase and number of scans per phase and is in keeping with low rates of new or worsening lesions in all but the Switch phase when the majority of patients were off DMTs, albeit for a relatively short period. **Figure 3-14** illustrates that this trend remains even when only (acute) contrast-enhancing lesions are considered, further confirming that the Switch phase was the period with the highest inflammatory radiological disease activity. Whilst the number of patients having MRI brain scans within the Switch period is lower than the other phases, potentially suggesting bias towards those with more active disease, it is notable that these MRI results largely mirror the ARR and SAD/SRD profiles seen in **Figures 3-7** and **3-8** with the Switch phase being an outlier of worsening disease activity and disability in comparison to other phases. Given this, however, the 'MRI efficacy cohort' with more complete clinical and radiological data is also considered below.

Table 3-21: Interval MRI Brain lesions by Treatment Phase – Full MRI cohort

N=36	Pre-Natalizumab	Natalizumab Treatment	Switch	Alemtuzumab Treatment	Post-Alemtuzumab
Number of patients	11 (31%)	16 (44%)	10 (28%)	15 (42%)	17 (47%)
Number of scans	19	71	13	25	48
Number of Gd scans	18	49	13	19	30
Mean duration of phase (years)	3.4	2.54	0.61	2	2.22
Worsening lesions					
Mean New Gd+ lesions (SD)	1.39 (2.1)	1.41 (4.5)	8.75 (9.7)	3.0 (9.6)	1.41 (6.0)
Mean New T2 lesions (SD)	3.21 (4.1)	1.30 (5.4)	4.53 (10.1)	1.35 (2.6)	2.04 (5.3)
Mean Worse T2 lesions (SD)	0.79 (1.3)	1.0 (2.0)	5.23 (2.4)	0.87 (1.8)	0.58 (2.5)
Mean New FLAIR lesions (SD)	3.63 (5.6)	1.47 (3.8)	5.75 (10.5)	1.65 (3.5)	1.86 (4.9)
Mean Worse FLAIR lesions (SD)	0.95 (1.8)	1.43 (3.6)	1.75 (2.5)	0.96 (2.1)	0.56 (2.3)
Mean New T1 lesions (SD)	1 (2.2)	0.65 (2.5)	2.91 (7.7)	0.63 (1.4)	1.25 (3.7)
Mean total any new or worse lesions*	9.58	5.85	20.17	5.46	6.29
Total new or worse lesions*/scan	0.50	0.08	1.55	0.22	0.13
Total new or worse lesions*/scan/year	0.15	0.03	2.54	0.11	0.06
Improving lesions					

Mean Improved T2 lesions (SD)	0.79 (1.2)	1.41 (4.1)	2.69 (6.0)	2.91 (5.9)	0.85 (2.9)
Mean Improved FLAIR lesions (SD)	0.68 (1.1)	1.64 (4.4)	3.08 (6.6)	3.83 (7.5)	1.35 (5.1)
Mean total any improved lesions*	1.47	3.05	5.77	6.74	2.2
Total improved lesions*/scan	0.08	0.04	0.44	0.27	0.05
Total improved lesions*/scan/year	0.02	0.02	0.73	0.13	0.02
Net effect (Worse-Improving) lesions/MRI/year	**0.13**	**0.01**	**1.81**	**-0.02**	**0.04**

*Excluding Gd+ lesions (denominator is total number of scans in phase)

Figure 3-11: New/Worse MRI Brain lesions by type and treatment phase – Full MRI cohort

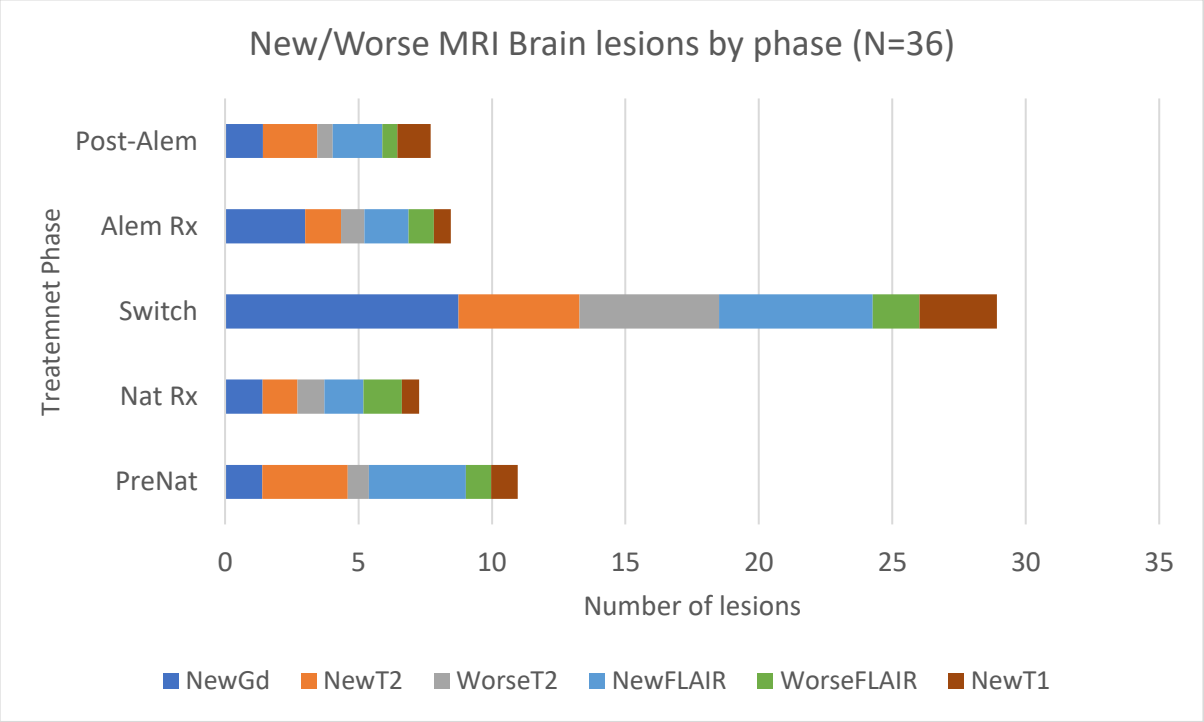


Figure 3-12: Improved MRI Brain lesions by type and treatment phase – Full MRI cohort

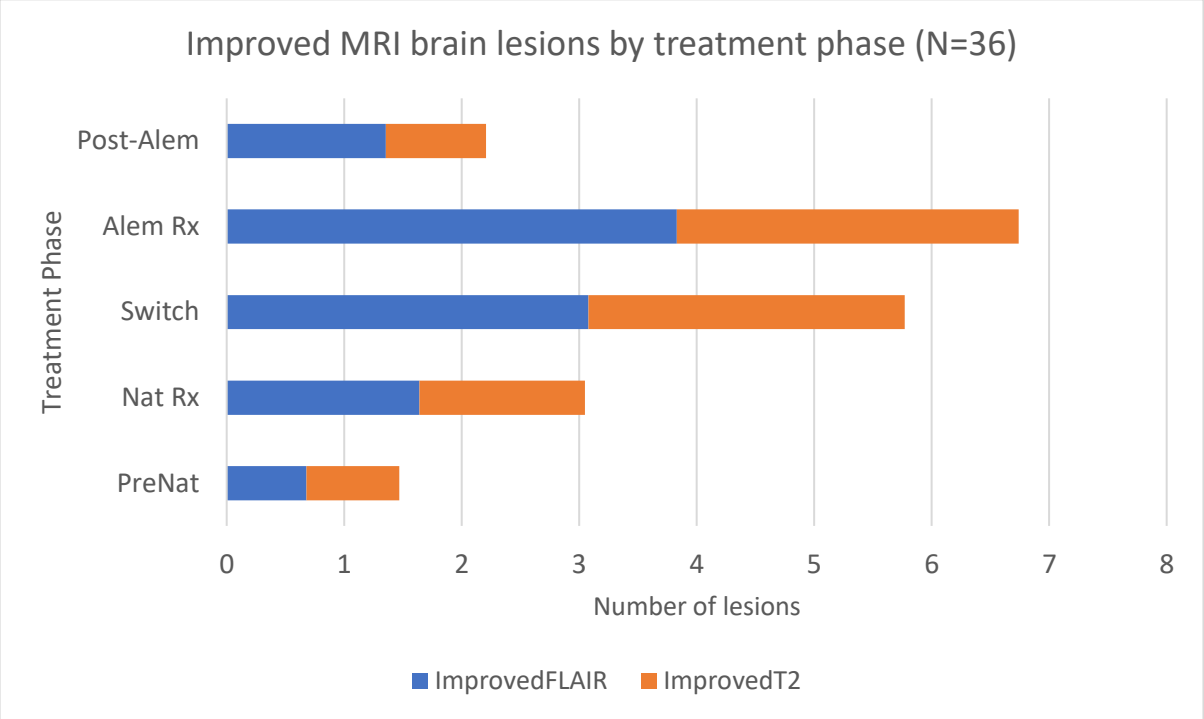


Figure 3-13: Total New/Worse MRI Brain lesions by treatment phase – Full MRI cohort

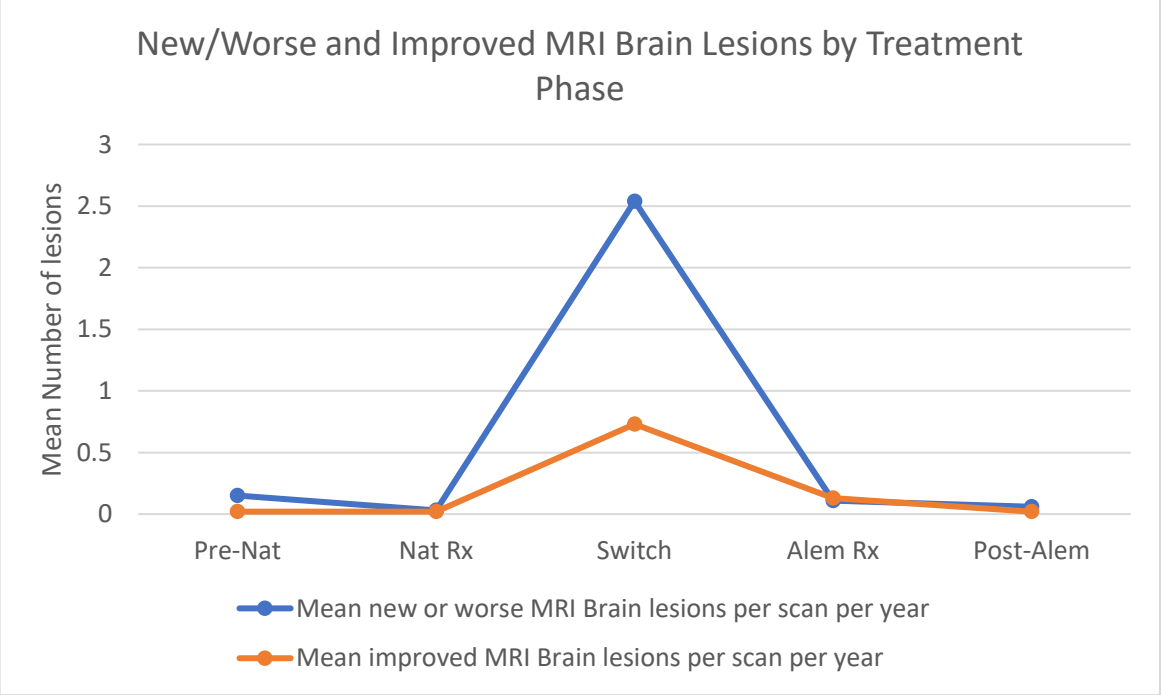
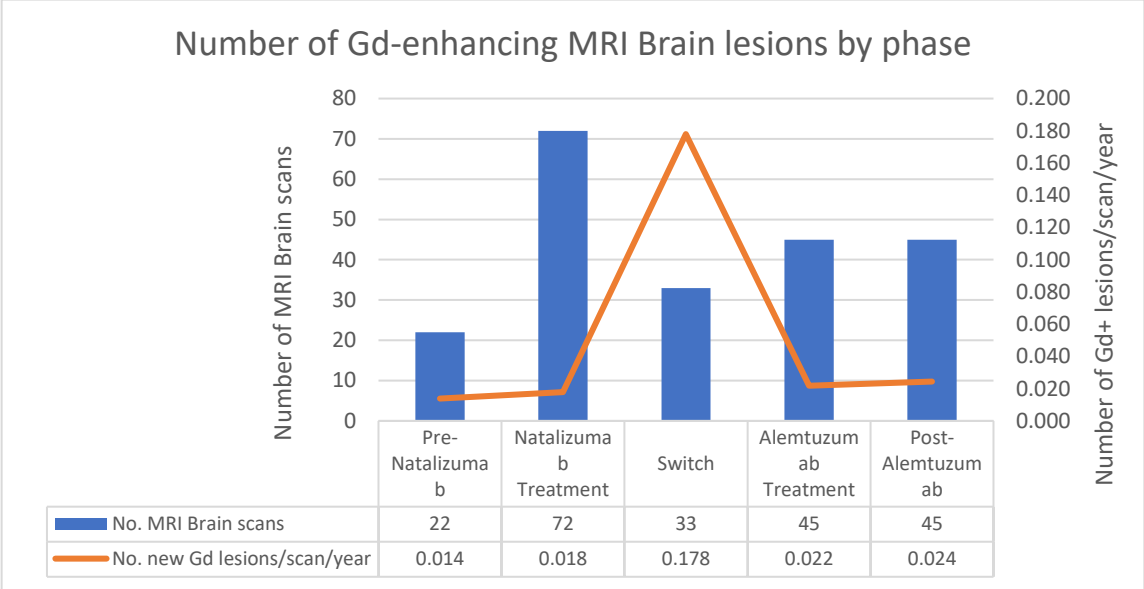


Figure 3-14: Number of contrast-enhancing lesions in each treatment phase



In addition to MRI Brain imaging, serial spinal cord imaging was assessed by a blinded neuroradiologist over the duration of the study where available. Overall, 25 patients from 6 centres had at least one MRI Cervical spine comparison during a treatment phase, totalling 56 MRI Cervical Spine comparisons over the duration of the study. Again, MRI disease activity was lowest in the Natalizumab, Alemtuzumab and Post-alemtuzumab phases but the number of patients and scans in each phase is small (See **Table 3-22**). **Figure 3-15** shows the total number of new T2 and contrast-enhancing lesions per treatment phase and in **Figure 3-16** this is adjusted for the number of scans undertaken and duration of each treatment phase. This corrects for the longer Pre-natalizumab phase and, as in previous results, disease activity is highest during the Switch phase, but the low number of new lesions in the (longest) Post-alemtuzumab phase suggests a persistent effect not seen in other efficacy measures in this study. These data largely concur with the previous evidence of high inflammatory disease activity within the Switch phase, low during other treatment phases, particularly Natalizumab and Alemtuzumab treatment, but also suggest a lasting effect of alemtuzumab not seen in other results. The small number of patients and scans limits any specific conclusions other than appearing to concur with the previous efficacy results for MRI brain imaging and relapses.

Table 3-22: MRI Cervical Spine lesions by treatment phase

	Pre-Natalizumab	Natalizumab treatment	Switch Period	Alemtuzumab Treatment	Post-Alemtuzumab
Total patients	18	5	3	5	8
Total MRI scans	18	15	3	6	14
Total MRI scans with contrast	14	15	3	6	6
Mean duration of phase [years] (SD)	3.5(2.9)	2.5(2.2)	0.6 (0.6)	2 (0)	4.3 (2.6)
Mean number of new T2 lesions (SD)	2.7 (2.3)	0.7 (1.0)	1 (1)	0.3 (0.8)	0.1 (0.3)
Mean number of new Gd+ lesions	0.5 (1.2)	0.3 (0.5)	0.3 (0.6)	0.2 (0.4)	0
Mean number of new T2 lesions /scan/year	0.04	0.02	0.6	0.03	0.002

Figure 3-15: MRI Cervical Spine Total new lesions by treatment phase

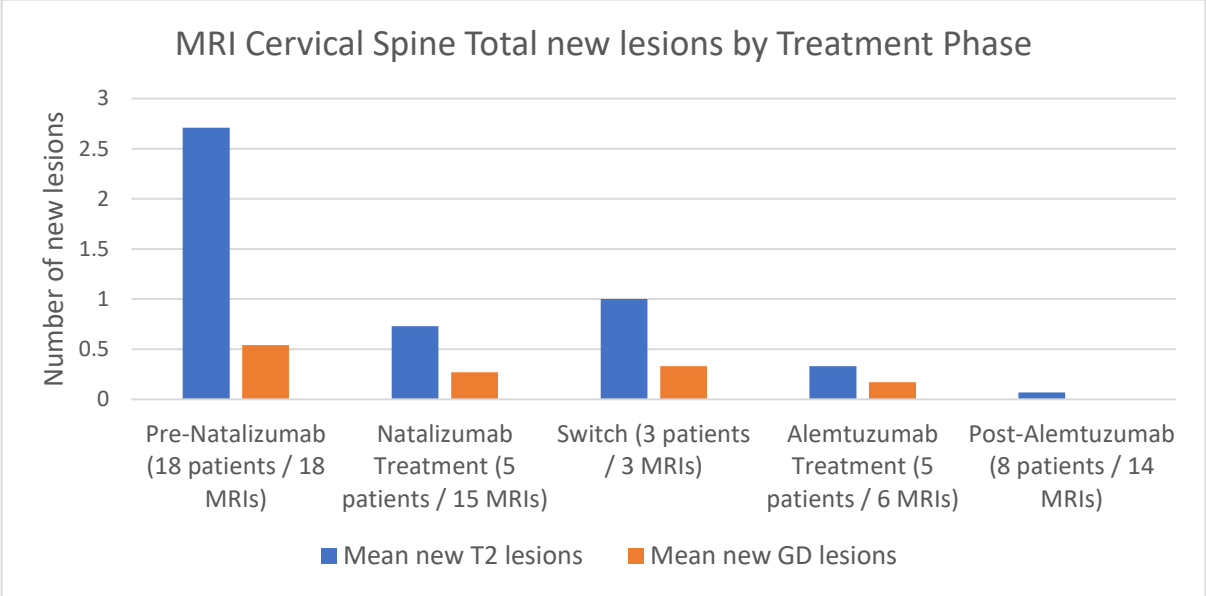
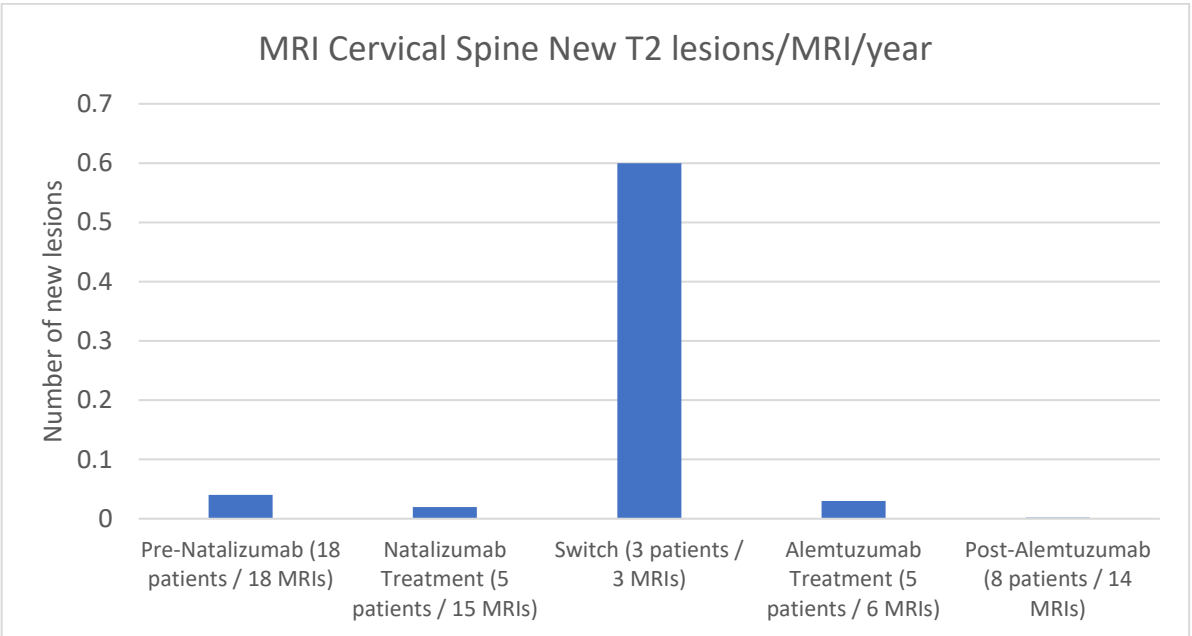


Figure 3-16: MRI Cervical Spine New T2 lesions / MRI / year of treatment phase



Specific MRI Brain comparisons were undertaken in an attempt to evaluate natalizumab and alemtuzumab treatment effects on lesion load and atrophy over time. The blinded neuroradiologist was provided with dates of scans for comparison generated by the online database and requested to compare specific

scans to assess potential treatment effects. Four comparison were made where available:

- The Latest MRI Brain in the Natalizumab treatment phase was compared with the latest MRI scan in the Pre-natalizumab phase to evaluate the effect of natalizumab treatment (N=7) [Mean treatment duration =1.6 years]
- The latest MRI Brain in the Alemtuzumab treatment phase was compared with the latest MRI scan in the Switch phase to evaluate the effect of alemtuzumab treatment within the first 2 years (N=18) [Mean treatment duration = 2 years]
- The latest available MRI Brain in the Post-alemtuzumab phase (i.e. most recent available) was compared with the latest scan in the Alemtuzumab treatment phase to evaluate the effect of alemtuzumab after 2 years (N=12) [Mean duration = 3.4 years]
- The latest available MRI Brain in the Post-alemtuzumab phase (i.e. most recent available) was also compared with the latest available scan in the Switch period to evaluate the overall effect of alemtuzumab (N=15) [Mean duration =5.4 years]

Overall, the number of new MRI Brain lesions was lower with alemtuzumab treatment in comparison to natalizumab treatment but the number of improved lesions was higher with natalizumab (See **Table 3-23** and **Figures 3-17**, **3-18** and **3-19**). The lower number of new or worsening lesions with alemtuzumab was largely driven by effects beyond 2 years of treatment in this cohort, suggesting a delayed effect on radiological disease activity. Conversely, the rate of improved MRI brain lesions was highest within the first 2 years after alemtuzumab treatment. Again, the number of patients is small, particularly for natalizumab treatment.

Table 3-23: MRI Brain lesions per year of Natalizumab and Alemtuzumab Treatment phase

	Natalizumab Treatment	Alemtuzumab Treatment in first 2 years	Alemtuzumab Treatment after 2 years	Overall Alemtuzumab Treatment
N	7	18	12	15
Mean New Gd+	0.88	0.17	0.79	0.51
Mean New T2	4.0	3.21	1.55	1.91
Mean New FLAIR	2.31	3.36	1.43	1.83
Mean Worse T2	0.44	0.34	0.53	0.34
Mean Worse FLAIR	0.18	0.40	0.41	0.51
Total New/Worse	**7.81**	**7.48**	**4.71**	**5.1**
Mean Improved T2	2.14	1.98	0.18	0.86
Mean Improved FLAIR	1.35	2.23	0.39	0.80
Total Improved	**3.49**	**4.21**	**0.57**	**1.66**

Figure 3-17: MRI Brain Specific Comparisons per year of Treatment - Mean New/Worse lesions

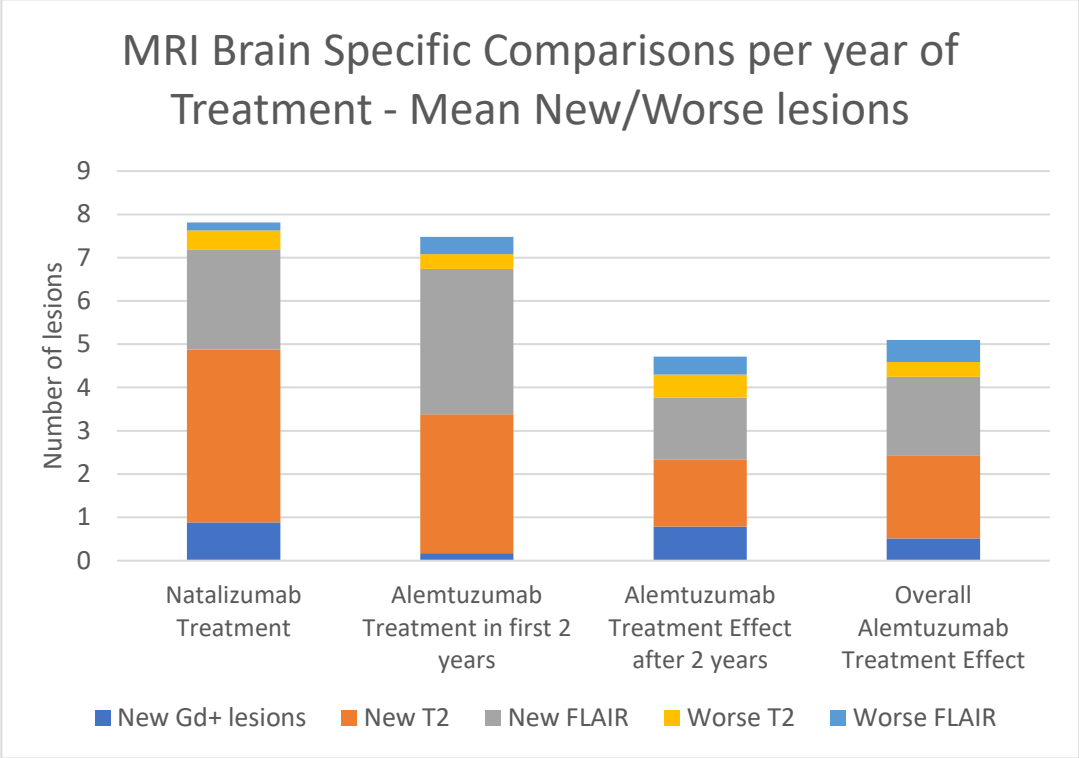


Figure 3-18: MRI Brain Specific Comparisons per year of Treatment - Mean Improved lesions

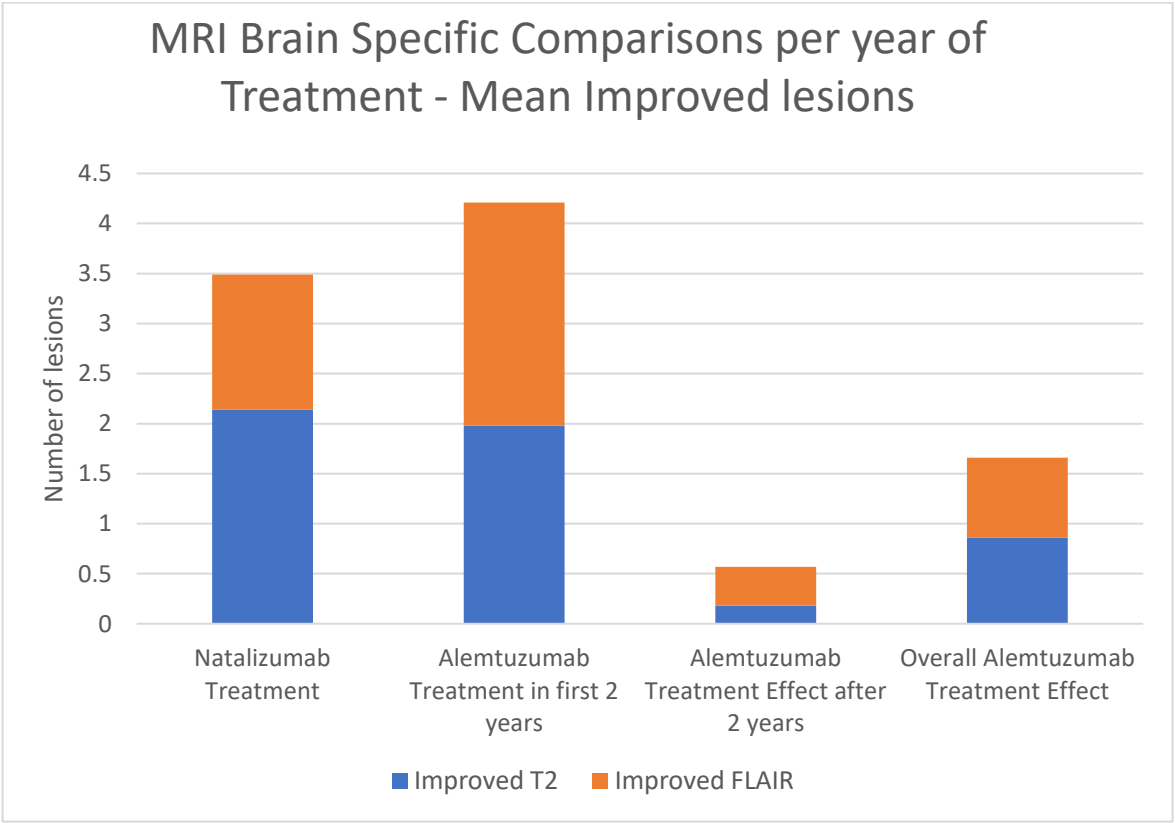
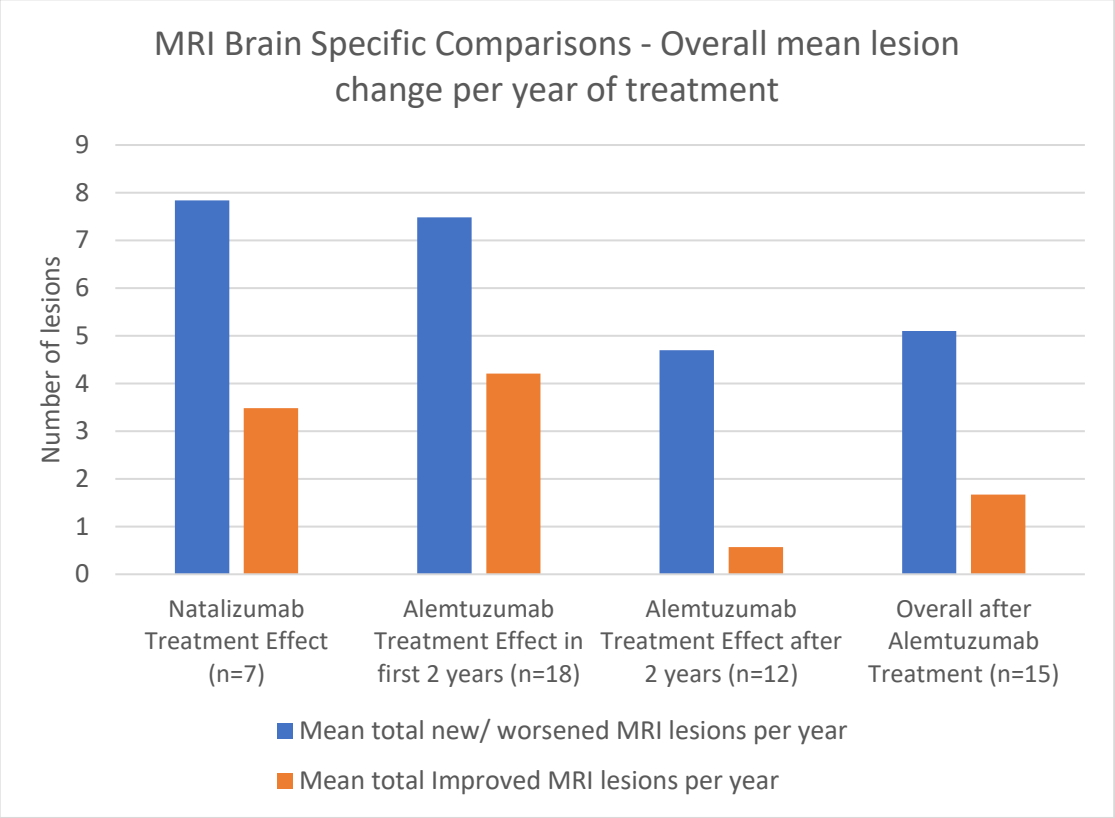
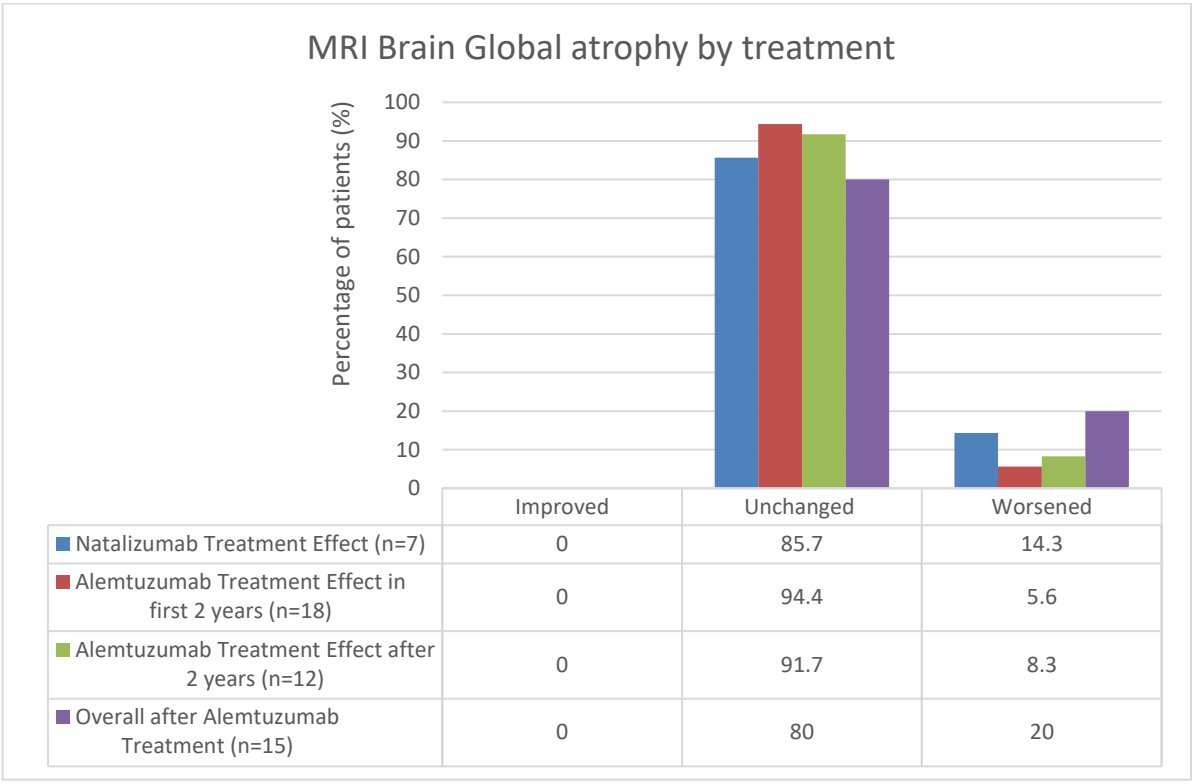


Figure 3-19: MRI Brain Specific Comparisons - Overall mean lesion change per year of treatment



In addition to changes in MRI brain lesion load, the blinded neuroradiologists also evaluated changes in brain atrophy over time during these treatment periods. As shown in **Figure 3-20**, the vast majority of patients had no visible change in brain atrophy by visual inspection with both natalizumab and alemtuzumab. Callosal atrophy was also assessed and the results mirrored that of global atrophy. No patients were deemed to have evidence of improved atrophy. Overall, brain atrophy was present in a higher proportion of patients during alemtuzumab treatment in comparison to natalizumab treatment. The cohort was older by the time of alemtuzumab treatment and had longer follow-up which may be relevant to the apparently worse atrophy during alemtuzumab treatment but, on average, these patients would still be in their late 30s by the end of follow-up making this unlikely to be a significant factor. The difference is small and the method non-systematic, but this suggests ongoing brain atrophy over time despite sequential treatment with highly effective DMTs.

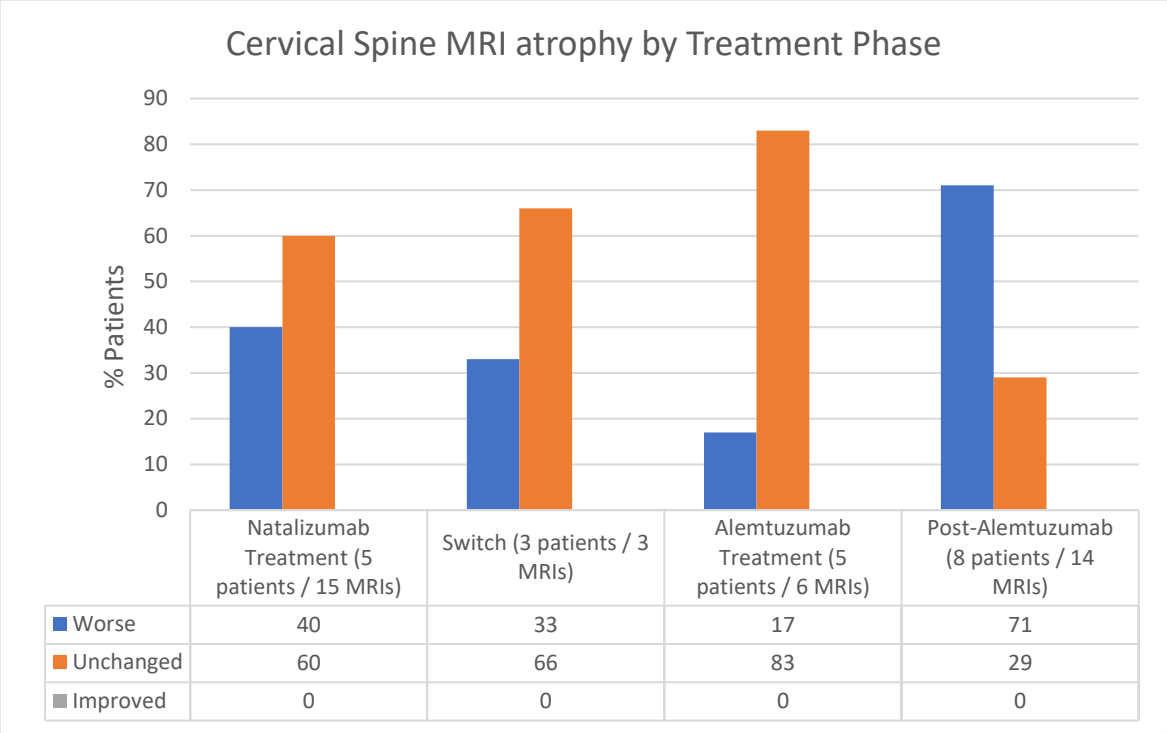
Figure 3-20: Global brain MRI atrophy changes with Natalizumab and Alemtuzumab



	Improved	Unchanged	Worsened
Natalizumab Treatment Effect (n=7)	0	85.7	14.3
Alemtuzumab Treatment Effect in first 2 years (n=18)	0	94.4	5.6
Alemtuzumab Treatment Effect after 2 years (n=12)	0	91.7	8.3
Overall after Alemtuzumab Treatment (n=15)	0	80	20

Cervical spinal cord atrophy was visually assessed across all treatment phases in patients with available MRI scans. Unlike the above analysis, this was across all treatment phases in the study rather than only during the Natalizumab and Alemtuzumab treatment phases but, again, the vast majority of patients had no evidence of change using this method (See **Figure 3-21**). As seen with brain atrophy evaluations, no patient had evidence of improvement in cervical cord atrophy. In contrast to the MRI brain atrophy findings, however, cervical spine atrophy was more commonly seen during natalizumab treatment than in the first 2 years of alemtuzumab treatment but the highest proportion of patients with cervical cord atrophy was reported during the Post-alemtuzumab phase. This again may suggest a treatment effect of alemtuzumab within the first 2 years which is not sustained in the longer term in this cohort.

Figure 3-21: Cervical Spine MRI atrophy by treatment phase



	Natalizumab Treatment (5 patients / 15 MRIs)	Switch (3 patients / 3 MRIs)	Alemtuzumab Treatment (5 patients / 6 MRIs)	Post-Alemtuzumab (8 patients / 14 MRIs)
Worse	40	33	17	71
Unchanged	60	66	83	29
Improved	0	0	0	0

Complete MRI and Clinical Data Cohort (N=25)

From the efficacy cohort, 25 patients (49%) had both clinical and MRI data available, with follow-up greater than 2 years from first alemtuzumab treatment, and these are analysed as a subgroup here in order to evaluate MRI outcomes in a well-defined clinical population (See **Table 3-2**). The characteristics of the Baseline MRI brain i.e. first available are outlined in **Table 3-24** and are largely in keeping with the total MRI cohort. The MRI efficacy cohort had a higher proportion of females than the total MRI cohort and were slightly older at diagnosis. Additionally, on average, disease duration at diagnosis was shorter in the MRI efficacy cohort but baseline EDSS and ARR were comparable. Fewer patients in the MRI efficacy cohort had natalizumab treatment as a 1st- line DMT and mean ARR and EDSS was lower in the Pre-natalizumab phase in comparison to the total MRI cohort, albeit the duration was the same. The MRI efficacy cohort had less supratentorial lesion load and atrophy in comparison to the total cohort but more infratentorial and contrast-enhancing lesions in the Pre-natalizumab phase. Treatment with natalizumab

was comparable but the Switch phase was shorter in the MRI efficacy cohort and they had longer available follow-up after alemtuzumab on average (5.3 vs 4.2 years). The MRI efficacy cohort started alemtuzumab at a younger age (33 vs 35) and with a shorter disease duration (6.7 vs 7.5) in comparison to the total MRI cohort but over a third required further DMTs after alemtuzumab treatment in both cohorts. Overall, the MRI efficacy cohort is comparable to the Total MRI cohort on most clinical measures but there is some suggestion of relatively milder disease at onset but more active disease despite DMTs prior to alemtuzumab.

The findings of the sequential interval MRI lesions change over time in this subgroup of patients with both clinical and MRI data available are largely in keeping with Total MRI cohort (See **Table 3-25** and **Figures 3-22** and **3-23**). Again, the Switch period is associated with the highest rate of new and worsening lesions in comparison to the other phases. In this cohort, however, the Alemtuzumab treatment phase is associated with the lowest rate of new or worsening MRI brain lesions and the Natalizumab treatment phase with the most improving lesions, having been the opposite in the Total MRI cohort. The magnitude of the radiological disease activity is also higher in this cohort in comparison to the Total MRI cohort, with a particularly significant reduction in the Alemtuzumab phase, following very high rates of activity during the Switch.

Using a two-tailed t-test, the rate of new or worsening lesions per MRI per year was significantly lower in the Alemtuzumab treatment period in comparison to the Switch period ($p=0.013$), Natalizumab treatment period ($p<0.001$) and Post-alemtuzumab period ($p<0.001$). There was no significant difference in new or worsening lesions between the Post-alemtuzumab and Natalizumab treatment periods however ($p=0.054$) but, similar to the Total MRI cohort, the improvements in inflammatory radiological disease activity during the Alemtuzumab treatment phase were not sustained in the Post-alemtuzumab phase. Indeed, the net effects are much the same per phase as described previously save for the fact that all groups had net worsening of their lesion load whereas alemtuzumab treatment associated with a (small) net benefit in radiological disease activity in the Total MRI cohort.

Table 3-24: Baseline MRI Brain scans – MRI Efficacy cohort

Baseline MRI (N=20)				
Mean number of Gd+ lesions* (SD, range)	6.9 (12.4, 0-38)			
Number of patients with no Gd+ lesions* (%)	6 (55%)			
Lesion load qualitative assessment	None	Minor	Moderate	Marked
Supratentorial	0	3	2	15
Infratentorial	4	4	2	7
Atrophy qualitative assessment				
Global	13	7	2	1
Callosal	12	5	2	1

*n=11 (MRIs where contrast was given)

Table 3-25: Interval MRI Brain lesions by Treatment Phase – MRI Efficacy cohort

	Pre-Natalizumab	Natalizumab Treatment	Switch	Alemtuzumab Treatment	Post-Alemtuzumab
Number of patients	8 (32%)	9 (36%)	4 (16%)	7 (28%)	16 (64%)
Number of scans	13	31	4	10	39
Number of Gd scans	12	25	4	7	28
Mean duration of phase (years)	3.4	1.97	0.37	2	2.94
Worsening lesions					
Mean New Gd+ lesions (SD)	1.46 (2.3)	1.44 (4.5)	8.75 (17.5)	0	2.23 (7.6)
Mean New T2 lesions (SD)	2.43 (3.7)	1.68 (5.3)	11.25 (18.1)	0.3 (0.7)	2.69 (7.5)
Mean Worse T2 lesions (SD)	0.36 (0.8)	0.58 (2.0)	0.25 (0.5)	0	1.13 (4.1)
Mean New FLAIR lesions (SD)	2.29 (4.5)	1.27 (3.8)	10.25 (18.6)	0.2 (0.4)	2.51 (7.2)
Mean Worse FLAIR lesions (SD)	0.21(0.8)	1.0 (3.6)	0.25 (0.5)	0	1 (3.5)
Mean New T1 lesions (SD)	1.31 (2.5)	0.9 (2.5)	7.5 (13.1)	0.2 (0.4)	1.64 (4.6)
Mean total any new or worse lesions*	8.06	6.87	38.25	0.7	11.2
Total new or worse lesions*/scan	0.62	0.22	9.56	0.07	0.29
Total new or worse lesions*/scan/year	0.18	0.11	25.84	0.04	0.10
Improving lesions					
Mean Improved T2 lesions (SD)	0.71 (1.1)	1.23 (4.2)	4.75 (8.9)	0.2 (0.4)	1.13 (3.3)
Mean Improved FLAIR lesions (SD)	0.57 (0.9)	1.23 (4.4)	5.0 (9.4)	0.4 (1.0)	2.05 (6.2)
Mean total any improved lesions*	1.28	2.46	9.75	0.6	3.18
Total improved lesions*/scan	0.10	0.08	2.43	0.06	0.08
Total improved lesions*/scan/year	0.03	0.04	6.59	0.03	0.03
Net effect (Worsening - Improving) lesions/ MRI/year	**0.15**	**0.07**	**19.25**	**0.01**	**0.07**

*Excluding Gd+ lesions (denominator is total number of scans in phase)

Figure 3-22: New/Worse MRI Brain lesions by Treatment Phase – MRI Efficacy cohort (N=25)

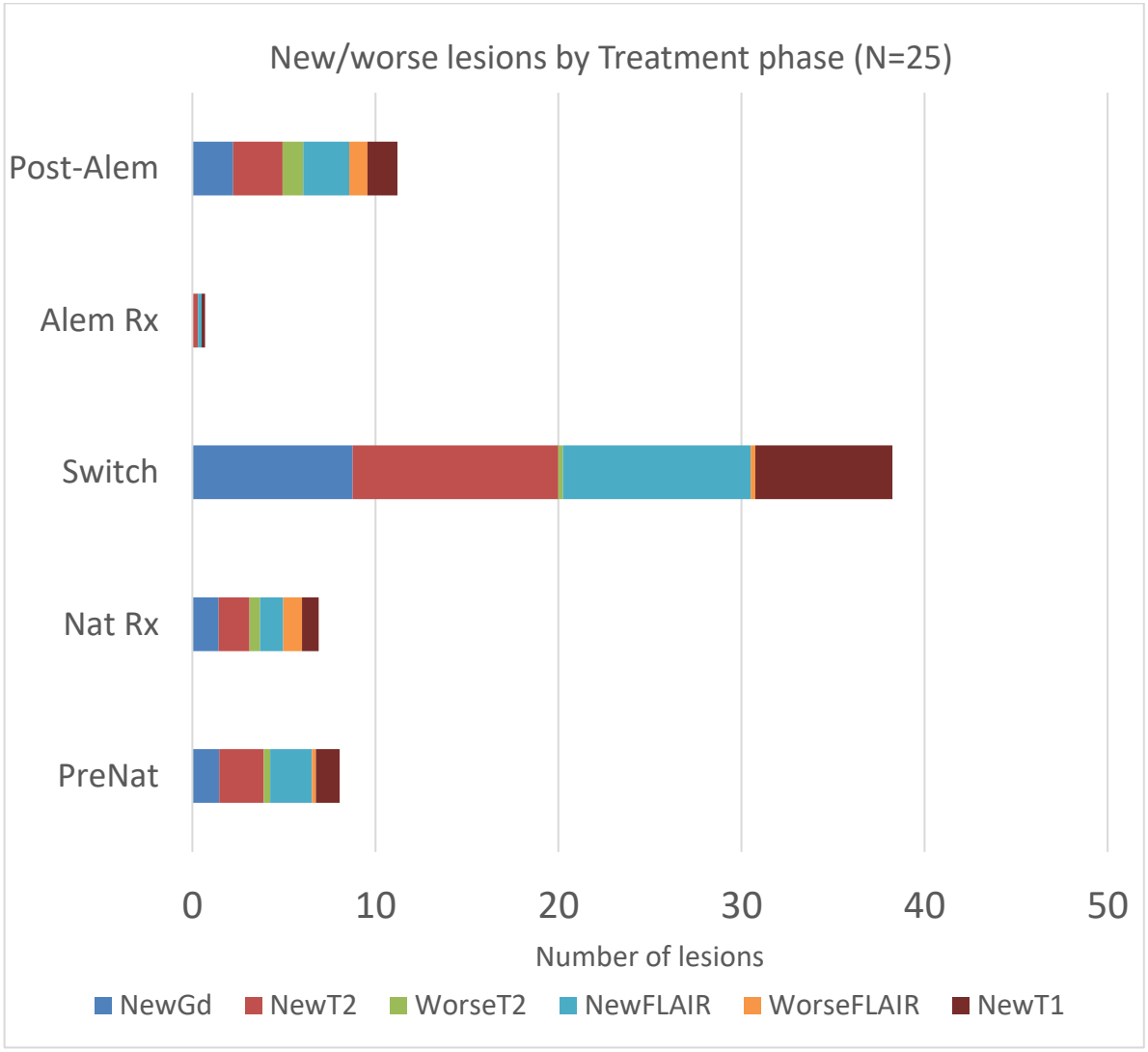
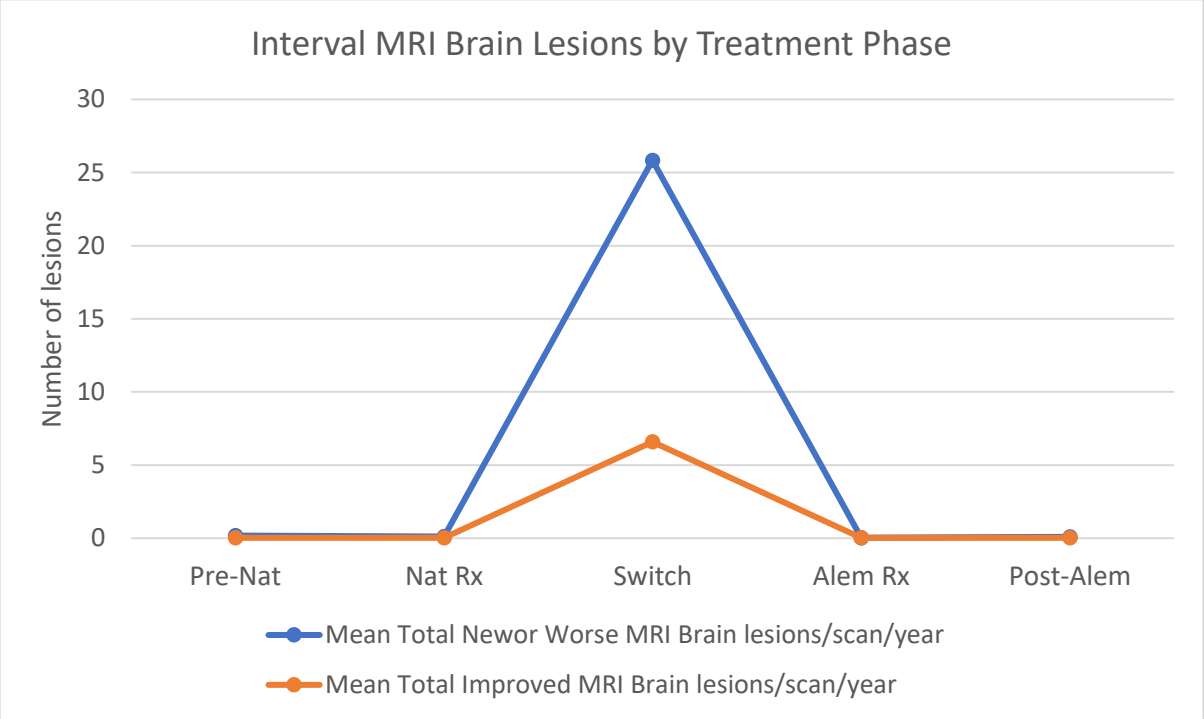


Figure 3-23: MRI Brain lesions/scan/year by Treatment Phase – MRI Efficacy cohort (N=25)



No evidence of Disease Activity (NEDA) by phase

The following Primary endpoint is addressed in this section:

- 'Freedom from disease activity' in each treatment phase

NEDA was defined as no relapses, increase in EDSS or MRI activity (T1/T2/FLAIR) ('NEDA-3'). This required patients to have data for all 3 modalities in each phase and, as outlined above, was the case for a minority only. Indeed, no patients had appropriate serial interval MRI Brain comparisons in every phase to allow sequential individual evaluation of lesion changes over the entire study. The data presented here are therefore for those patients in each phase where all three modalities were available but are not the same patients over time. In total, 19 patients had 98 interval comparison MRI brain scans available over the course of the study. **Figure 3-24** shows the proportion of patients in each phase who had MRI brain interval comparisons and no evidence of new disease activity.

Most of the patients with MRI data available also had relapse and EDSS data for the relevant phase but this was not the case for all and further reduces the denominator of cases to evaluate NEDA in each phase. **Table 3-26** outlines the overall proportion of patients with all necessary data available and the proportion with NEDA in each phase. The small numbers again limit any conclusions based on these data alone but, again in keeping with previous results, disease activity appears lowest in the Natalizumab, Alemtuzumab and Post-alemtuzumab phases. This is in spite of the variable treatment phase durations, with disease activity occurring even during the short Switch phase. Each patient had undergone MRI brain more than once in most cases, however, suggesting that these were clinically active or deteriorating to merit imaging. This is likely to mean selection bias with more active patients having one or more MRI scans but, even with this, some met criteria for NEDA-3 and the trends are in keeping with better disease control during natalizumab and alemtuzumab treatment.

Figure 3-24: Proportion of patients with no new MRI brain lesions by treatment phase

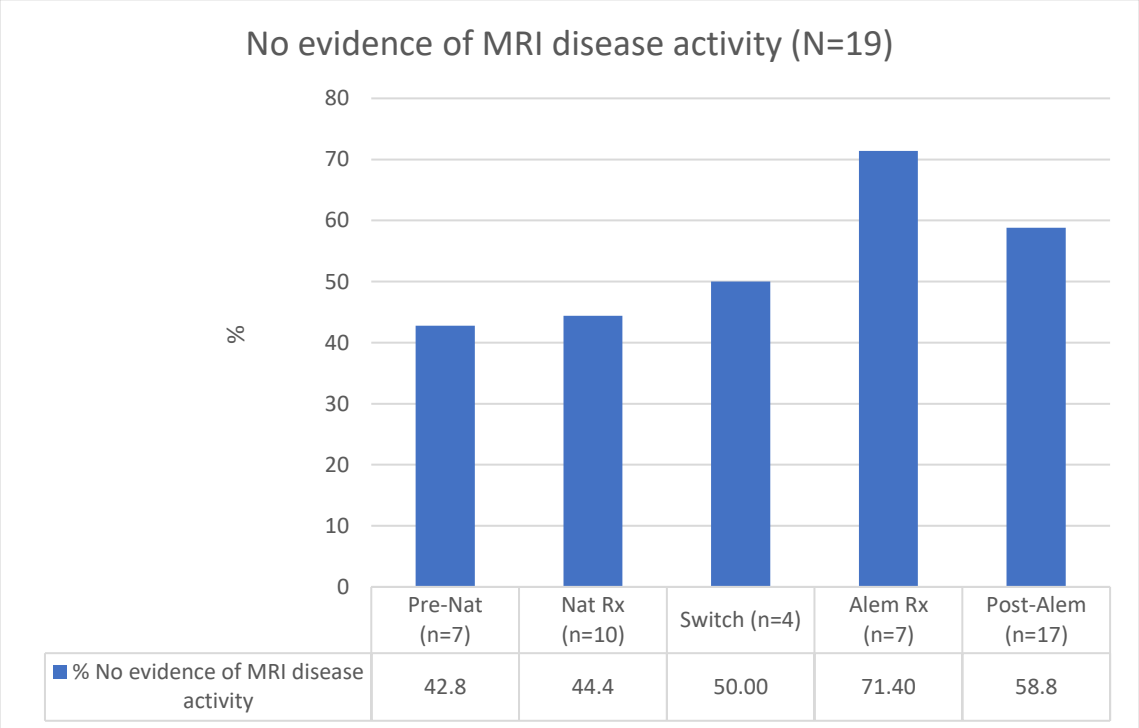


Table 3-26: NEDA-3 by phase (N=19)

	Pre-Natalizumab	Natalizumab Treatment	Switch	Alemtuzumab Rx	Post-alemtuzumab
Total number of patients	7	10	4	7	15
Total number of MRI Brain scans	14	31	4	10	39
Patients with NEDA-MRI	3 (42.8%)	5 (50%)	2 (50%)	5 (71.4%)	10 (58.8%)
Patients with NEDA-MRI and no Relapses	0	4	1	5	6
Patients with NEDA-MRI and no EDSS increase	0	3	0	3	3
Patients with NEDA-3 (%)	**0**	**3 (30%)**	**0**	**3 (42.8%)**	**1 (6.7%)**

EDSS Results

The following Secondary endpoint is addressed in this section:

- EDSS change during each treatment phase

In the whole cohort, mean EDSS scores increased up until the end of the Switch phase and then fell after alemtuzumab treatment, but missing data was an issue hence focus on the efficacy group here (N=51). **Table 3-27** shows data for the entire cohort and outlines the limited number of available EDSS scores in each phase, particularly at diagnosis (Baseline) and in the Post-alemtuzumab phase. Having anticipated the availability of approximately 6-monthly EDSS scores, analysis of the total cohort was not possible due to the lack of available data for a significant proportion and further analysis on those with the most available data is therefore the focus below. With these caveats, **Figure 3-25** charts the median EDSS and ARR for the whole cohort in each treatment phase and demonstrates a reduction in both following alemtuzumab treatment, albeit from their highest level during the Switch phase.

The trends are the same for efficacy cohort (N=51). The frequency of documented EDSS scores was higher in this group but the proportion that occurred during relapse or were estimated were comparable to that in the entire cohort. In contrast to the whole cohort, however, the median EDSS increases during the Alemtuzumab treatment phase and falls more modestly in the Post-alemtuzumab phase, more in keeping with a plateau rather than the reduction seen in the whole cohort. These data are represented in **Table 3-28** and **Figure 3-26**.

Table 3-27: EDSS by treatment phase (N=79)

N=79	Baseline	Pre-Natalizumab	Natalizumab Treatment	Switch	Alemtuzumab Treatment	Post-Alemtuzumab
Mean	2.4	3.4	4.3	4.7	4.4	4.3
Min	0	0	0	0	0	0
Max	8	8.5	8.5	8.5	8.5	8
Median	2	3	4.5	5.5	5	4
Number of patients with data [at least 1 EDSS] (%)	55 (69.6)	62 (78.4)	68 (86.1)	49 (62.0)	58 (73.4)	36 (45.6)
Number of documented EDSS scores	55	141	181	65	115	83
	Duration of Phase (days)					
	n=65					
Mean	6316	1555	876	235	730	599
Min	4	14	59	28	730	-458*
Max	41611	8052	3201	1149	730	2421
Median	192	1159	669	155	730	277
Total patient-days	404251	122875	69242	18538	57670	47309
Mean frequency of EDSS recorded (EDSSs / year)	0.05	0.42	0.96	1.3	0.73	0.64
Number (%) EDSS scores during relapse	15 (27%)	31 (22%)	23 (13%)	15 (23%)	8 (7%)	5 (6%)
Number (%) EDSS scores estimated	31 (56%)	21 (15%)	30 (17%)	14 (22%)	19 (17%)	16 (19%)
Dates of recorded EDSS scores	Nov 1988 – Nov 2016	March 2000 – Nov 2014	Jan 2007 – June 2015	March 2009 – Jan 2017	Oct 2009 – Nov 2017	Nov 2011 – Nov 2017

*Not all patients had a Post-Alemtuzumab phase i.e. <2yrs since Alemtuzumab first started at time of database closure for analysis

Figure 3-25: EDSS and ARR by treatment phase for whole cohort (N=79)

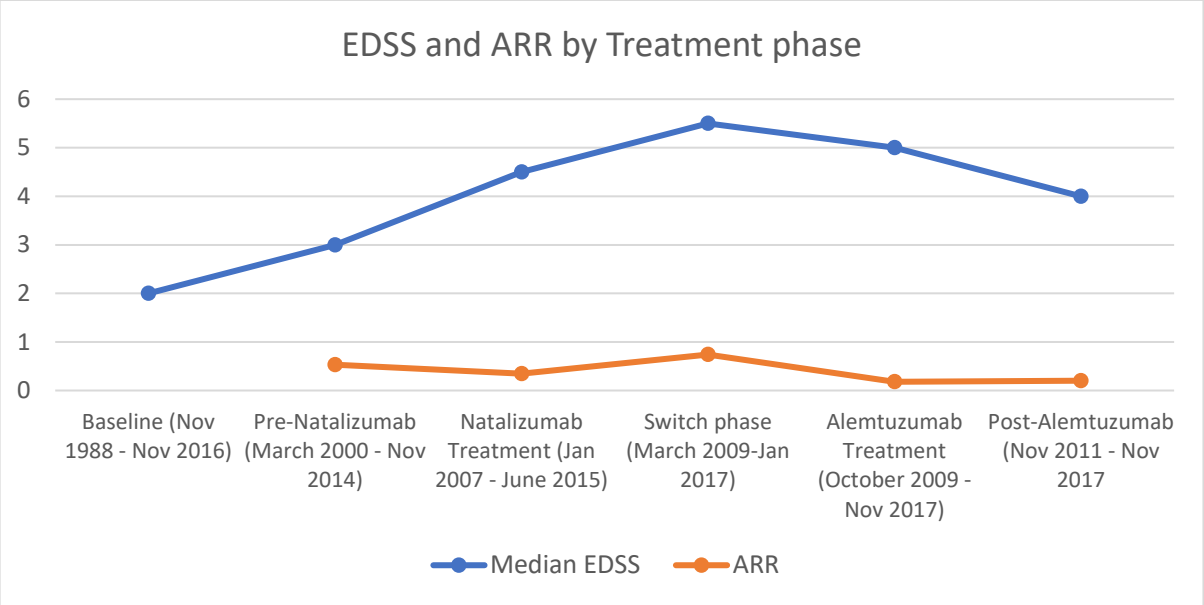
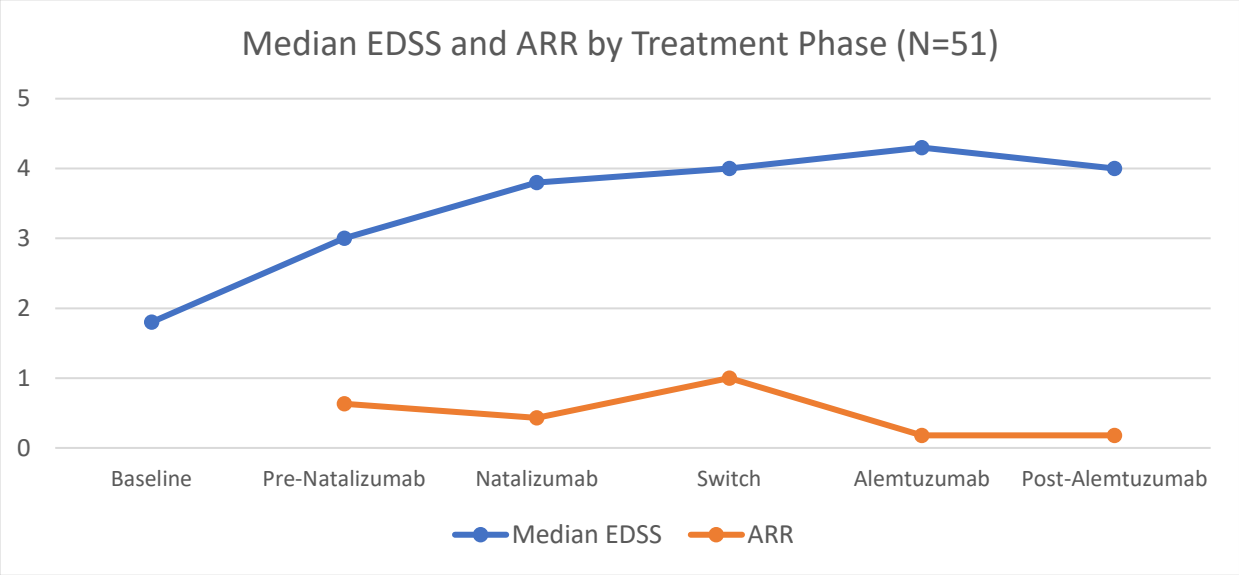


Table 3-28: EDSS by treatment phase [Efficacy cohort (N=51)]

N=51	Baseline	Pre-Natalizumab	Natalizumab Treatment	Switch	Alemtuzumab Treatment	Post-Alemtuzumab
Mean	2.4	3.2	3.9	4.1	4.2	4.2
Min	0	0	0	0	0	0
Max	8	8.5	8.0	8.0	7.5	8.0
Median	1.8	3	3.8	4.0	4.3	4.0
Number of patients with data [at least 1 EDSS] (%)	44 (86.3)	45 (88.2)	45 (88.2)	31 (60.8)	46 (90.2)	35 (68.6)
Number of documented EDSS scores	44	105	96	39	96	81
	Duration of Phase (days)					
	n=41					
Mean	309	1541	666	174	730	1023
Min	4	14	60	28	730	31
Max	3439	8052	2805	796	730	2421
Median	125	1211	423	117	730	1040
Total patient-days	12677	78601	33987	8854	37230	52196
Mean frequency of EDSS recorded	1.27	0.49	1.03	1.61	0.94	0.57

(EDSSs / year)						
Number (%) EDSS scores during relapse	11 (25%)	23 (22%)	18 (19%)	9 (23%)	5 (5%)	5 (6%)
Number (%) EDSS scores estimated	26 (59%)	16 (15%)	19 (20%)	8 (21%)	16 (17%)	16 (20%)
Dates of recorded EDSS scores	Nov 1988 - March 2014	March 2000 - Feb 2014	Jan 2007 - March 2015	March 2009 - October 2015	October 2009 - July 2017	November 2011 - November 2017

Figure 3-26: EDSS and ARR by treatment phase for Efficacy cohort (N=51)



Total study EDSS-years Area Under the Curve (AUC)

The AUC method was used to assess EDSS change over the whole study and in each phase, comparing changes in comparison to the initial baseline EDSS at diagnosis. The area under the curve (AUC) was calculated and divided by the number of years in the phase to give an annualised baseline-adjusted AUC for each patient in each phase where there was sufficient data. Negative values indicate better outcomes i.e. reduced disability.

A total of 35 patients had sufficient EDSS data to calculate their AUC across the entire study period. These patients had a median time in the study of 11 years

and a total of 356 documented EDSS scores, of which 12% were estimated and 13% occurred during a relapse.

Overall, the median annualised baseline-adjusted EDSS AUC was 1.26 EDSS-years. Over the median time of 11 years this would give a total AUC of 13.5 EDSS-years. There are many profiles that would give this total, for example an increase of 1 at the start followed by an increase of 0.5 for the last five years. In terms of a gradual continuous increase over the study period however, this would equate to an increase in EDSS of around 0.25 per year for each patient on average.

As would be expected in an active cohort, the vast majority of patients had a worsening AUC profile over the course of the whole study from diagnosis until last follow-up after alemtuzumab i.e. worsening disability. Only 4/35 (11%) improved overall compared to their baseline EDSS and a further 6/35 (17%) were unchanged but 71% of patients worsened over the study period (See **Table 3-29**).

Table 3-29: Overall Annualised Total AUC (baseline adjusted) and AUC categorisation

Total study (N=35)	**Time in study (yrs)**	**Max EDSS per patient**	**Annualised Total AUC (baseline adjusted)**
Mean±SD	11.73±4.96	5.46±2.28	1.21±1.95
Median±IQR	10.71±4.62	6.5±4.50	1.26±2.06
Range (min,max)	(4.84,29.16)	(2,8.5)	(-5.39,4.61)
		Net improved (%)	4 (11%)
		Net unchanged (%)	6 (17%)
		Net Worse (%)	25 (71%)
Number of documented EDSS scores		356	
Number (%) in relapse		47 (13%)	
Number (%) estimated		44 (12%)	

Notably, there was no correlation between annualised total AUC and time in study. Those with longer disease duration were no more likely to have a different rate of deterioration.

EDSS AUC by Treatment Phase

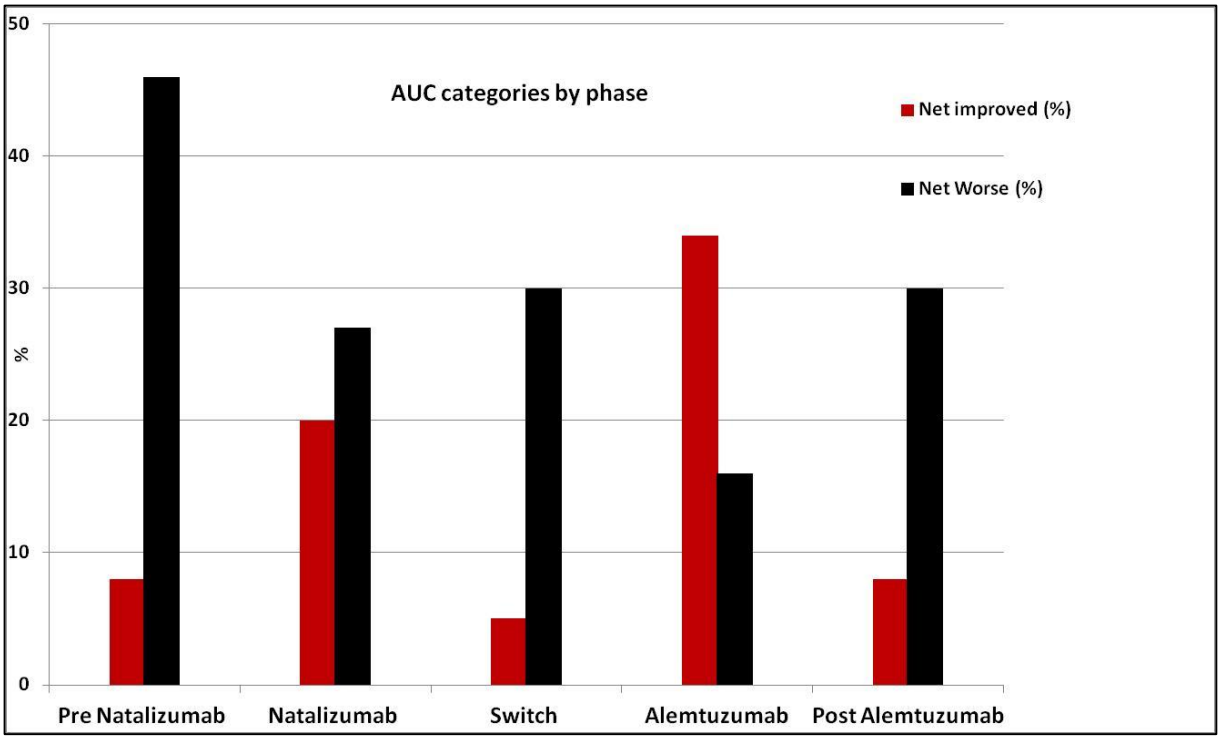
In the efficacy cohort (N=51), available EDSS scores were plotted against time in each phase for each patient with the vertical axis scale starting at the baseline EDSS score. Overall, mean annualised EDSS AUC fell from the Pre-natalizumab phase (0.5 EDSS-years) to Natalizumab treatment (0.2), rose slightly during the Switch phase (0.23), fell considerably during the Alemtuzumab treatment phase (-0.39) but rose again to almost Pre-natalizumab levels during the post-Alemtuzumab phase (0.45) (See **Table 3-30**).

Mean annualised EDSS AUC was further categorised as 'net improved' (for decreases of more than 0.5), 'net worse' (for increases of more than 0.5) and 'net unchanged' (for values in between). The proportion of 'net worse' decreased after the Pre-natalizumab stage but then remained similar for all phases except during the Alemtuzumab treatment phase, were there was a fall. The proportion of 'net improved' remained similar in the Pre-natalizumab, Switch and Post-alemtuzumab phases but increased in both the Natalizumab and Alemtuzumab treatment phases (See **Table 3-30** and **Figure 3-27**).

Table 3-30: EDSS and AUC analysis by phase (N=51)

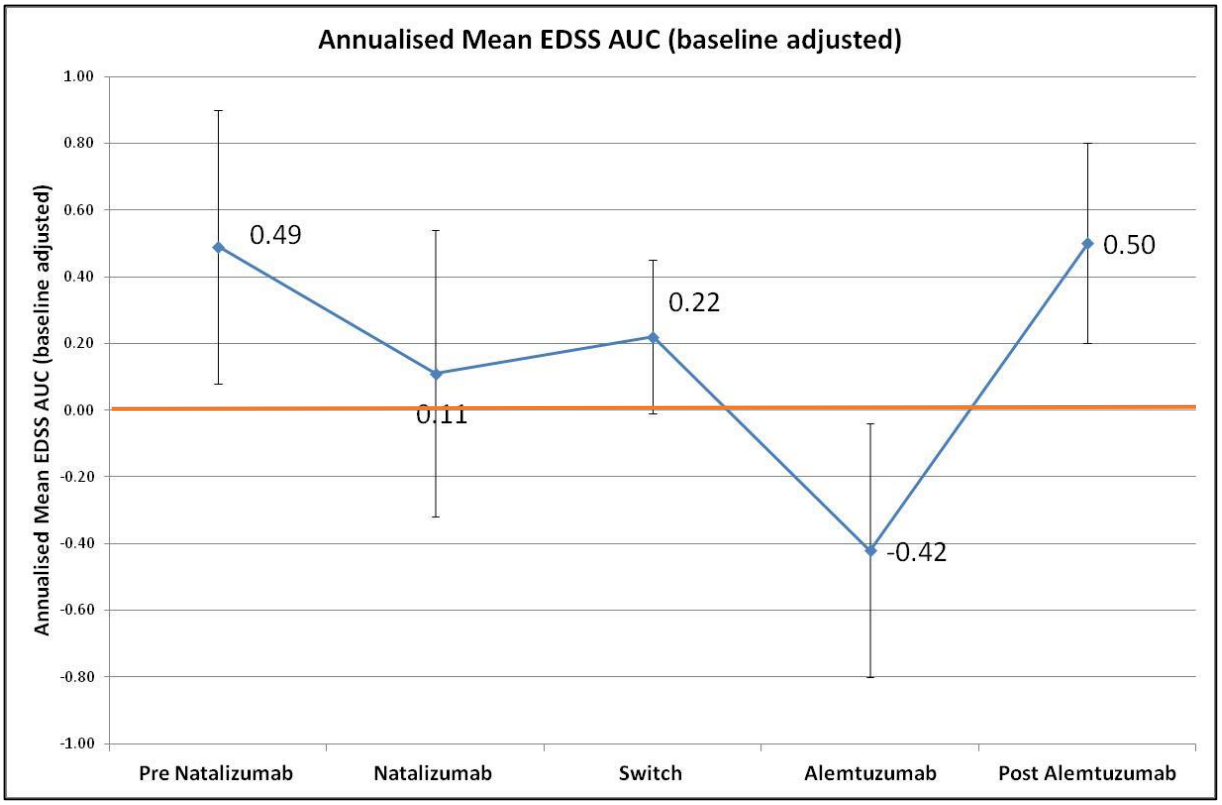
Phase (N)	Baseline (44)	Pre Natalizumab (45)	Natalizumab (47)	Switch (38)	Alemtuzumab (46)	Post Alemtuzumab (38)
Maximum EDSS per patient						
Mean (SD)	2.32 (1.71)	3.68 (2.03)	4.39 (2.23)	4.38 (2.17)	4.37 (2.22)	4.64 (2.22)
Median (IQR)	1.5 (1.5)	3.5 (3.25)	4.50 (4.5)	4.0 (4.1)	4.5 (4.5)	5.0 (4.1)
Range (min,max)	0,8	1,8.5	0,8	0,8	0,7.5	0,8
Annualised Total AUC (baseline adjusted)		(39)	(45)	(37)	(38)	(37)
Mean (SD)		0.50 (1.42)	0.20 (1.26)	0.23 (0.64)	-0.39 (1.07)	0.45 (0.89)
Median (IQR)		0.44 (1.49)	0 (0.64)	0 (0.53)	-0.11 (1.20)	0 (0.71)
Range (min,max)		-4.29,3.64	-3.98,3.22	-1.90,1.99	-3.90,1.37	-0.95,2.55
AUC categorisation						
Net improved (%)		3 (8%)	9 (20%)	2 (5%)	13 (34%)	3 (8%)
Net unchanged (%)		18 (46%)	24 (53%)	24 (65%)	19 (50%)	23 (62%)
Net Worse (%)		18 (46%)	12 (27%)	11 (30%)	6 (16%)	11 (30%)

Figure 3-27: EDSS AUC categorisation by phase (only improved or worse shown)



There were 35 patients with data for EDSS AUC analysis across all five phases [mean (SD) for each phase: 0.49 (1.19), 0.11 (1.25), 0.22 (0.66), -0.42 (1.10), 0.50 (0.87)]. A repeated measures ANOVA found that there was a significant fall in mean EDSS AUC between Switch and Alemtuzumab treatment phases (p=0.008) and a significant rise between Alemtuzumab treatment and Post - alemtuzumab (p=0.002). The changes between Pre-natalizumab and Natalizumab treatment and Natalizumab treatment and Switch were not significant (p=0.25 & 0.72). Non-parametric tests confirmed these results. **Figure 3-28** shows the mean AUC (& 95% error bars) across phases for this cohort, with the line of no change (0) highlighted: values above this reflect net disability worsening and values below improvement. Improvement in disability, on average, during the Alemtuzumab treatment phase was not maintained in the Post-alemtuzumab phase, where there was significant worsening.

Fig 3-28: Mean annualised EDSS AUC (baseline adjusted) by phase (N=35)



36 patients had EDSS AUC data for both the Natalizumab and Alemtuzumab treatment phases. Of these 14/36 (39%) were in the same category for both phases (i.e. net better/worse or unchanged suggesting no differing effect of either treatment) but 13/36 (36%) were better off in the Alemtuzumab treatment phase and 9/36 (25%) were worse off. However, there was no statistically significant difference between the Natalizumab and Alemtuzumab phases.

Overall, the EDSS AUC analysis is in keeping with our other results in that outcomes are generally best during the Natalizumab and Alemtuzumab treatment phases. In this analysis, the first 2-year period after alemtuzumab treatment was associated with the lowest proprtion of disability worsening of all phases and also the highest proprtion of patients with reduced disability. Direct comparison of the Natalizumab and Alemtuzumab treatment periods did not find a statistically signficant difference in AUC EDSS change however. The highly encouraging results obtained during the first 2 years after alemtuzumab

treatment were not maintained in the longer term however and, indeed, appear to show a return toward Pre-natalizumab rates of disability.

EDSS vs. SAD/SRD analysis

We have used two methods to analyse disability over time based on EDSS, namely SAD/SRD analysis and the EDSS-years AUC method. In comparison of these methods in our cohort, there was some agreement between the SAD/SRD profile categories and the total AUC categories. However, there were 10 patients who were recorded as 'confirmed stable' by SAD/SRD but were 'net worse' in the AUC analysis (See **Table 3-31**).

Table 3-31: SAD/SRD categorisation by AUC categorisation

SAD/SRD categories	Total AUC categories			Total
	Net improved	Net unchanged	Net worse	
Confirmed stable	4	6	10	20
Unsustained change	0	0	6	6
Confirmed worsening	0	0	9	9
Total	4	6	25	35

If the AUC criteria were changed so that less than zero (any negative AUC) was 'net improved' (rather than decreases of >0.5) and values between 0 and 1 was 'net unchanged' (rather than changes of <0.5), and only AUC above 1 was classified as 'net worse' (rather than increase >0.5) then there is slightly better agreement, with only 8 patients misclassified but this does not change the overall categorisation of net changes over the study, with the majority still 'net worse' (See **Table 3-32**).

Table 3-32: SAD/SRD categorisation with refined AUC categorisation

SAD/SRD categories	Total AUC categories			Total
	Net improved	Net unchanged	Net worse	
Confirmed stable	5	7	8	20
Unsustained change	0	2	4	6
Confirmed worsening	0	0	9	9
Total	5	9	21	35

Overall, it is likely that the AUC method is most accurate in this study as the SAD/SRD analysis relies more on frequently measured EDSS scores over time.

Efficacy Results – Summary

The majority of the clinical and radiological efficacy results suggest the lowest levels of most disease activity markers in patients treated with alemtuzumab after natalizumab were during the first 2 years after alemtuzumab treatment, with lasting positive effects on inflammatory disease activity (relapses / some MRI lesions) but not disability or MRI atrophy. A summary of efficacy results are illustrtaed in **Table 3-33**, where the treatment phase with the best performance for the outcome is highlighted in green and the least favourable treatment phase in red; additionally, whether the beneficial effects of the 2-year Alemtuzumab treatment period are maintained into the Post-alemtuzumab phase (or not) are highlighted similarly.

The highest levels of disease activity across most domains were during the Switch period between natalizumab and alemtuzumab. The outcomes during the Alemtuzumab and Natalizumab treatment phases were largely comparable in most domains, save for MRI atrophy measures and disability worsening using the AUC method where the early Alemtuzumab treatment period had the more positive effects.

The small numbers make conclusions less than definitive but there appears to be agreement throughout the majority of clinical and radiological outcome measures used here that patients failing natalizumab treatment in this cohort have the highest levels of disease activity/worsening during the Switch phase

but good response to alemtuzumab treatment initially, with sustained reductions in some measures of inflammatory disease activity, but benefits on disability and MRI atrophy are not maintained. The dissociation between inflammatory and 'neurodegenerative' measures of disease in MS are well-described although it seems this is only the case in this cohort once the 2-year treatment phase of alemtuzumab is completed.

Table 3-33: Clinicoradiological Efficacy Results Summary (N=51 unless otherwise stated)*

	Pre-Natalizumab	Natalizumab Treatment	Switch Period	Alemtuzumab Treatment	Post-Alemtuzumab	Alemtuzumab Treatment effect maintained?
Clinical outcome measures						
ARR	0.57	0.43	1.1	0.18	0.18	Yes
Annualised SAD rate	7.0	6.1	11.2	5.7	6.6	No
Annualised SRD rate	1.8	6.1	5.6	10.2	0.8	No
Median EDSS	3	3.8	4.0	4.3	4.0	-
Mean EDSS AUC	0.5	0.2	0.23	-0.39	0.45	No
AUC Category Net improved Net Unchanged Net Worse	 8% 46% 46%	 20% 53% 27%	 5% 65% 30%	 34% 50% 16%	 8% 62% 30%	 No - No
MRI outcome measures						
Mean net new/worse MRI Brain lesions per MRI per year (N=36)	0.13	0.01	1.81	-0.02	0.04	No
Mean net new/worse MRI Brain lesions per MRI per year (N=25)	0.15	0.07	19.25	0.01	0.07	No
Mean new/worse MRI Brain lesions in specific MRI comparisons	-	7.81	-	7.48	4.71	Yes
Mean improved MRI Brain lesions in	-	3.49	-	4.21	0.57	No

specific MRI comparisons						
Number of new Gd+ lesions on MRI brain / scan / year	0.014	0.018	0.18	0.021	0.024	No
Mean new T2 MRI C-spine lesions	0.04	0.02	0.6	0.03	0.002	Yes
% Worse Global Atrophy in Specific comparisons	-	14.3	-	5.6	8.3	No
% Worse Cervical Spine atrophy	-	40	33	17	71	No
% No evidence of new MRI lesions	42.8	44.4	50	71.4	58.8	No
Combined outcome measure						
% NEDA-3	0	30	0	42	6.7	No

*The treatment phase with the best performance is highlighted in green, the lowest in red

DMT-related Neutralising Antibodies

The following Secondary Endpoint is addressed in this section:

- Presence of neutralizing antibodies in each treatment phase

Neutralising antibodies (Nabs) were only recorded only during the Natalizumab treatment phase (See **Table 3-2**). Natalizumab NAbs were documented in 7 of 79 patients (8.8%): 5 of these patients stopped natalizumab due to efficacy, 1 due to hypersensitivity reaction (SAE) and 1 because of NAbs alone. These data are in keeping with the known deleterious effect of NAbs on the efficacy of natalizumab.

Quality of Life markers

The following Secondary Endpoint is addressed in this section:

- Change in any marker of quality of life / employment

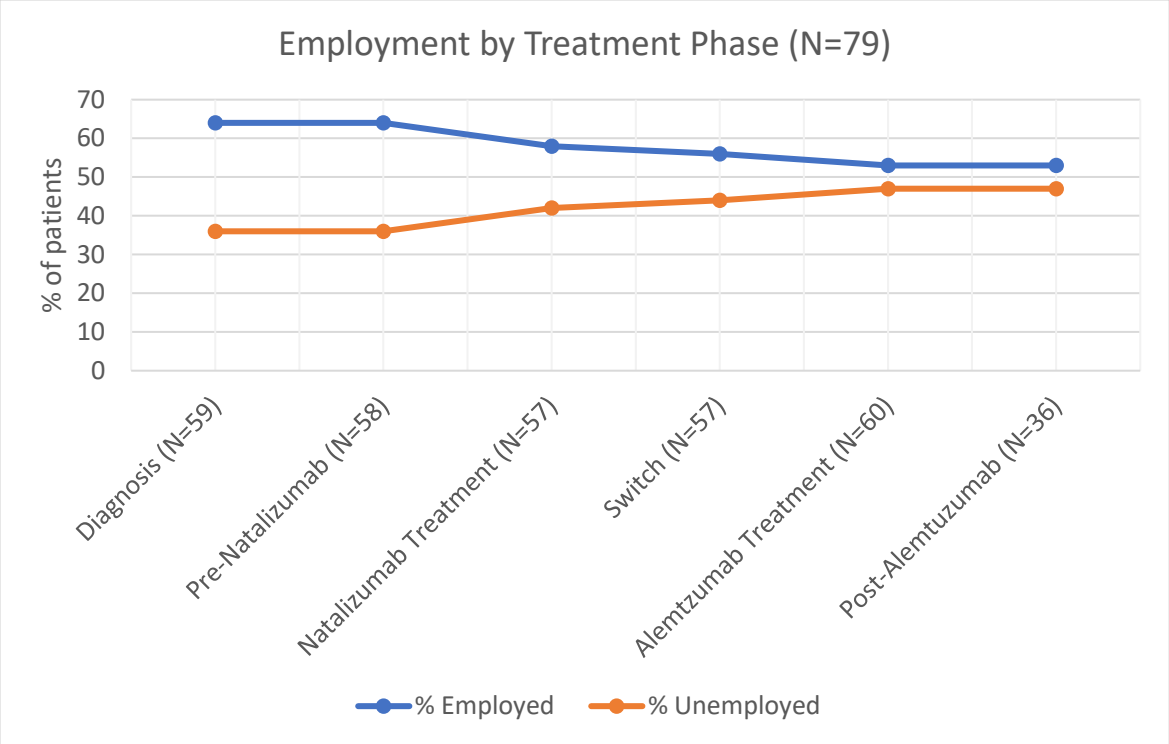
Employment

The majority of patients remained employed throughout each phase of the study. The proportion of patients in employment fell with time in the total cohort, however. At diagnosis, 64% were in employment but this had fallen to 53% by last follow-up in the Post-alemtuzumab phase, although the number of patients with available data had fallen (See **Table 3-34** and **Figure 3-29**). This was not just due to unavailable data however, as the proportion unemployed increased with time. Few patients became employed having been unemployed in the preceding treatment phase but this appeared to be most common between diagnosis and starting DMTs initially. Equally, a small proportion of patients became unemployed between phases but this was more consistent over time.

Table 3-34: Employment by Treatment Phase

N=79	Diagnosis	Pre-Natalizumab	Natalizumab Treatment	Switch	Alemtuzumab Treatment	Post-Alemtuzumab
N	59	58	57	57	60	36
Employed	38 (64%)	37 (64%)	33 (58%)	32 (56%)	32 (53%)	19 (53%)
Unemployed	21 (36%)	21 (36%)	24 (42%)	25 (44%)	28 (47%)	17 (47%)
Unemployed to Employed	0	4 (7%)	1 (2%)	1 (2%)	1 (2%)	1 (2%)
Employed to Unemployed	0	4 (7%)	4 (7%)	2 (4%)	3 (5%)	0

Figure 3-29: Employment by Treatment Phase



In the efficacy cohort, there were data on 40/51 patients (22% missing) throughout all phases. In keeping with the total cohort, the majority [26/40 (65%)] were employed at diagnosis. Over the study period 11/14 (79%) of those unemployed at diagnosis remained unemployed throughout, while 1 person (7%) became employed and remained so thereafter. Of those in employment at

diagnosis, 16/26 (62%) remained employed throughout. A large proportion of these [6/26 (23%)] became unemployed and did not regain employment thereafter, however. Some patients [6 (15%)] oscillated between employment and unemployment over the course of the study. The biggest change was from initial employment to unemployment overall, in keeping with results from the total cohort.

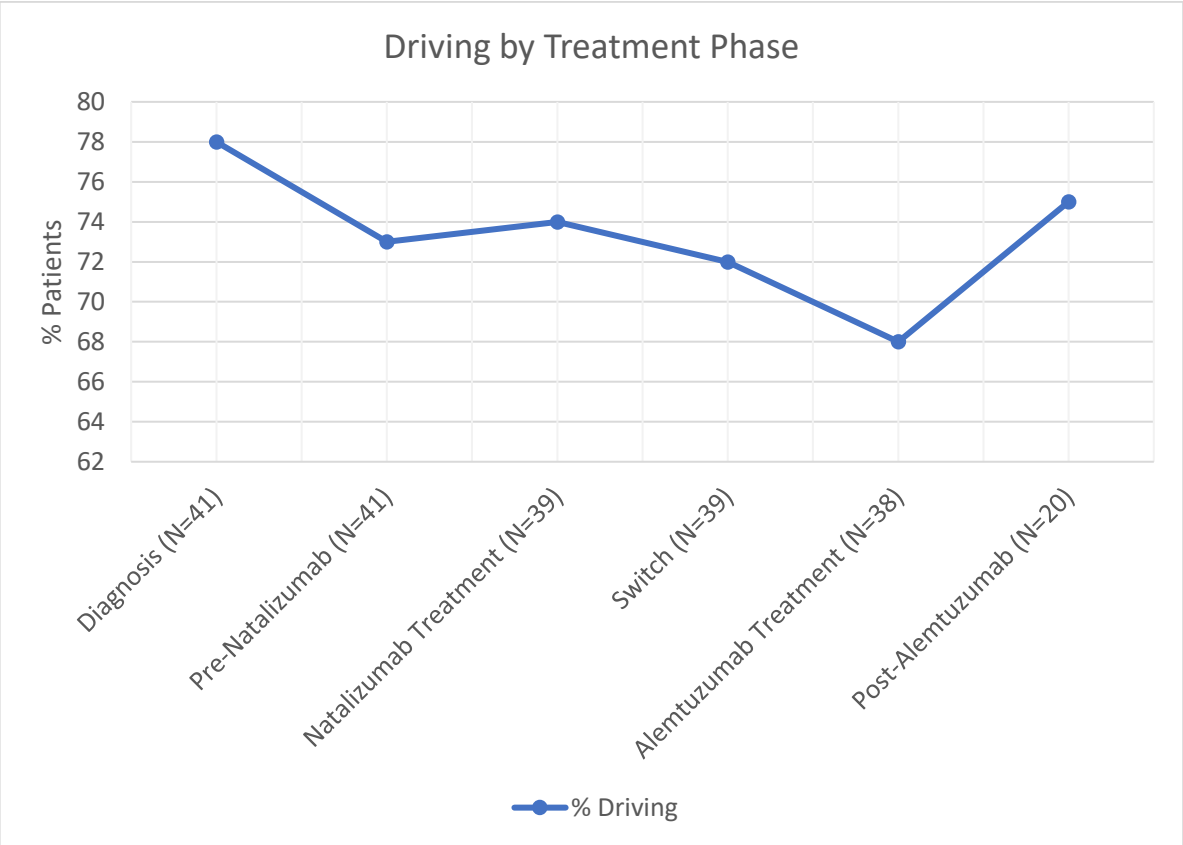
Driving

The majority of the cohort were drivers at diagnosis and remained so over time although almost half of patients in the total cohort did not have driving data available. Similar to employment data, the proportion of drivers fell over time although driving appeared to increase back towards baseline in the Post-alemtuzumab phase but the number of patients with data in this phase was again significantly lower. A small proportion of patients stopped driving in each treatment phase save for the Post-alemtuzumab phase. These data are illustrated in **Table 3-35** and **Figure 3-30**.

Table 3-35: Driving by Treatment phase

N=79	Diagnosis	Pre-Natalizumab	Natalizumab Treatment	Switch	Alemtuzumab Treatment	Post-Alemtuzumab
N	41	41	39	39	38	20
Driving	32 (78%)	33 (73%)	29 (74%)	28 (72%)	26 (68%)	15 (75%)
Not driving	9 (22%)	8 (27%)	10 (26%)	11 (28%)	12 (32%)	5 (25%)
Not driving to driving	0	1 (2%)	0	0	0	2 (10%)
Driving to not driving	0	1 (2%)	1 (2.5%)	1 (2.5%)	1 (3%)	0

Figure 3-30: Driving by Treatment phase



In the efficacy cohort, 21/51 patients (59% missing) had driving data available in all phases of the study. Again, the majority [16/21 (76%)] were driving at the time of diagnosis and 5/21 (24%) were not. Over the study period 12/16 (75%) continued driving while 2 patients stopped driving (one between diagnosis and the Pre-natalizumab phase, one after starting natalizumab). A further 2 patients oscillated between driving and not driving (one between Natalizumab treatment, the Switch period and Post-alemtuzumab; one between the Switch period, Alemtuzumab treatment and Post-alemtuzumab). Of those not driving at diagnosis, 3/5 (60%) remained not driving throughout but 2/5 (40%) started driving and continued thereafter (one between diagnosis and the Pre-natalizumab phase and one after Alemtuzumab treatment).

Deprivation category

A Deprivation Index score is available for all patients with a UK postcode (N=65) using the Index of Multiple Deprivation (IMD) from the national statistics offices of England, Scotland and Wales. This is provided with the decile of the relevant area, based on the Lower Layer Super Output Area (LSOA) code for patients in England and the postcode for patients in Scotland and Wales. Notably, only the Outward Code (Postcode Area and District) were collected to avoid patient-identifiable level data in the study. This therefore is not entirely accurate to small areas but the Outward Code was cross-referenced with the LSOA code and the deprivation centile based on this for England. **Table 3-36** lists the number of patients in each decile, where decile 1 is the most deprived and decile 10 the least. The majority of patients [44/65 (68%)] are in the lower half of deprivation deciles, suggesting a relatively deprived cohort overall.

Table 3-36: Deprivation Category Decile for UK cohort at diagnosis

Deprivation Category Index Decile	N
1 (most deprived)	7
2	9
3	8
4	9
5	11
6	11
7	3
8	2
9	2
10 (least deprived)	3
Total	65

Over the period under study only 4 patients changed postcode and one patient changed twice. The postcode change was from a lower decile (more deprived) area to a higher (less deprived) area for 2 patients although one of these patients then returned to the lower decile area. For these patients, the change occurred after alemtuzumab for one patient and after natalizumab treatment started for the other, who then returned to the lower decile area during the Post-alemtuzumab phase. Two patients changed from a higher decile area to lower decile areas, both occurring after the Alemtuzumab treatment phase.

Pregnancies

In total, 26 pregnancies were documented in the entire cohort over the course of the study. Only a single complication was recorded and this was outwith any DMTs: a retrospectively noted fetal heart condition in a previous pregnancy at the time of diagnosis. There was a large amount of missing data for this however. The number of pregnancies fell with time but so did the number of women with available data as shown in **Table 3-37**.

Table 3-37: Pregnancies by treatment phase

N=79	Pre-diagnosis	Pre-natalizumab	Natalizumab treatment	Switch	Alemtuzumab treatment	Post-Alemtuzumab
Total number of women with data	35	13	15	10	12	6
Number of pregnancies	16	3	4	0	2	1
Number of women	10	3	4	0	2	1
Complications	Fetal heart condition					

Symptomatic therapy

The use of symptomatic therapy was recorded in all phases and the results from the Efficacy Cohort are presented here (N=51). The highest number of symptomatic therapies were used in the Alemtuzumab & Post-alemtuzumab phases, which accounted for over half of all symptomatic treatments used over the study period [124/239 (52%)]. The other phases had smaller numbers. The difference between the Alemtuzumab phases and the other phases was significant (chi-square $p<0.01$).

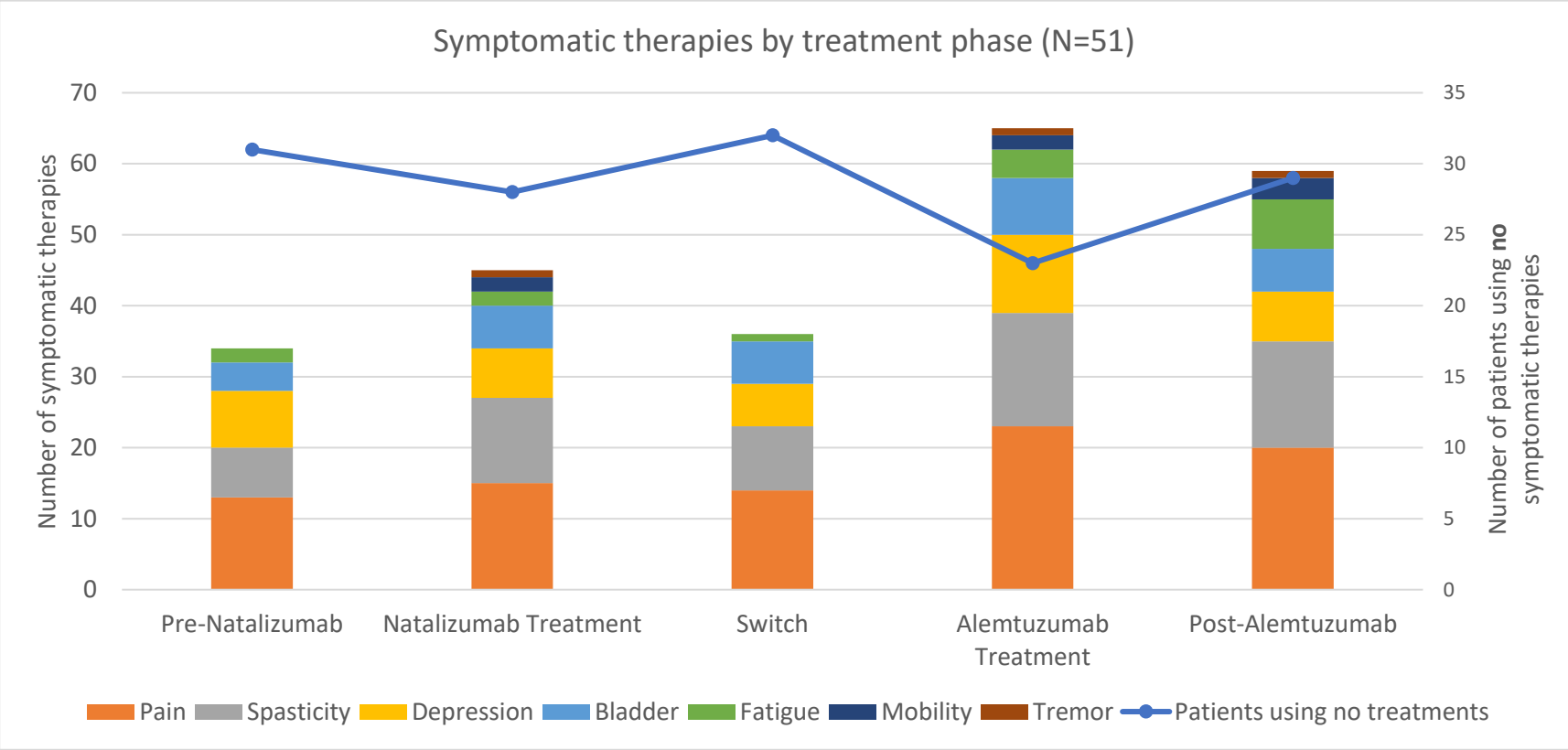
Overall, the symptoms most often treated were pain and spasticity, comprising 61% of all treatments (36% & 25% respectively). Depression and bladder problems accounted for another 29% (16% & 13%) while treatments for fatigue, mobility and tremor were less common accounting for only 11%. These category profiles were broadly similar in all phases. Whilst there were more treatments in the Alemtuzumab phases, the proportions for each treatment category was broadly in line with the other phases. **Table 3-38** and **Figure 3-31** detail these results.

Table 3-38: Symptomatic therapies by Treatment phase

N=51	Pre-Natalizumab	Natalizumab treatment	Switch	Alemtuzumab treatment	Post-Alemtuzumab	Overall total
Number of patients with no therapy (%) (95%CI†)	31 (61%) (47%,73%)	28 (55%) (41%,63%)	32 (63%) (49%,75%)	23 (45%) (32%,59%)	29 (57%) (43%,69%)	
Mean (SD)	0.67 (1.03)	0.88 (1.18)	0.71 (1.10)	1.27 (1.60)	1.16 (1.85)	
Total therapies (max per patient)	34 (4)	45 (4)	36 (4)	65 (4)	59 (8)	239
Categories						
Pain	13 (38%)	15 (33%)	14 (39%)	23 (35%)	20 (34%)	85 (36%)
Spasticity	7 (21%)	12 (27%)	9 (25%)	16 (25%)	15 (25%)	59 (25%)
Depression	8 (24%)	7 (16%)	6 (17%)	11 (17%)	7 (12%)	39 (16%)
Bladder problems	4 (12%)	6 (16%)	6 (17%)	8 (12%)	6 (10%)	30 13%)
Fatigue	2 (6%)	2 (4%)	1 (3%)	4 (6%)	7 (12%)	16 (7%)
Mobility	0	2 (4%)	0	2 (3%)	3 (5%)	7 (3%)
Tremor	0	1 (2%)	0	1 (2%)	1 (2%)	3 (1%)

†Using the Wilson Score method

Figure 3-31: Symptomatic therapies used by Treatment phase



Cognition and ambulation

Two further secondary endpoints were planned but could not be evaluated due to lack of data. No cognitive assessments were documented for any patient during the study and the number of ambulation scores was inadequate for any meaningful analysis. The occurrence of new clinical markers of cognitive impairment and changes in ambulation score are therefore not included.

The Switch Period

The following Secondary endpoint is addressed in this section:

- To evaluate whether there is an optimal time period for the 'switch' phase between natalizumab and alemtuzumab and the effect of alternative DMTs as bridging therapies

Cohort characteristics by length of switch period

In order to evaluate whether there is an optimal time period to switch from natalizumab to alemtuzumab, the Efficacy Cohort (N=51) was divided into those who switched in less than 4 months and those who took longer. The 4-month cut-off was to provide similar numbers in each group to evaluate subsequent outcomes but also has a basis from other studies suggesting improved outcomes in patients switching from natalizumab to fingolimod in less than 16 weeks[239]. The Switch period was defined as the time between the last infusion of natalizumab and the first infusion of alemtuzumab. The characteristics of those switching from natalizumab to alemtuzumab in less than or greater than 4 months (120 days) are outlined in **Table 3-39** in comparison to the total Efficacy Cohort.

There were some notable differences between the patients switching to alemtuzumab from natalizumab in less than or greater than 4 months (See **Table 3-39**). The difference in duration of the Switch period and the use of DMTs only in the longer switch group were statistically significant. Additionally, the higher proportion of JCV antibody positive patients in the longer switch group was

significantly different in comparison to those switching in less than 4 months, even though it appears the proportion in each group stopping natalizumab due to concern about PML was comparable but a quarter of the longer switch group did not have their reason for stopping natalizumab documented. The longer switch group also had shorter follow-up than those who switched in less than 4 months, meaning fewer outcomes will be captured in this study.

The relatively small number of patients in the study means that some differences may not be statistically significant due to underpowering but other potentially relevant differences are noted. The shorter switch group had more females and was older on average at diagnosis. They had a shorter disease duration at diagnosis on average and shorter duration between diagnosis and commencing natalizumab. Despite being older on average at diagnosis, they were younger by the time of both natalizumab and alemtuzumab treatment than those in the longer switch group. This may suggest an older population with more aggressive disease leading to earlier treatments. The higher baseline EDSS in the shorter switch group may also suggest this but they had had fewer pre-diagnosis relapses on average in comparison to the longer switch group. Also, the shorter switch group had fewer MRI Brain lesions present during the Pre-natalizumab phase and lower EDSS before natalizumab treatment. However, the relapse rate prior to natalizumab was higher in the shorter switch group and natalizumab was more likely to be used as a first-line therapy. The duration of natalizumab treatment (and number of infusions) was notably shorter in the shorter switch group and more of these patients failed natalizumab due to efficacy rather than PML concerns, which was the case for a quarter of patients in the longer switch group who also had the statistically significantly higher proportion being JCV positive. The shorter switch group were more likely to require further courses of alemtuzumab and other DMTs in the Post-alemtuzumab phase, again suggesting a relatively refractory treatment group. Overall, it seems those switching from natalizumab to alemtuzumab in less than 4 months had more active disease and there was more concern about the risk of PML with ongoing use of natalizumab in the longer switch group, both of which likely resulted in the shorter and longer switch periods respectively.

Notably, there appears also to have been influence on the duration of the switch period between natalizumab and alemtuzumab by the year in which it occurred. In the whole efficacy cohort, those who started natalizumab before 2010 (23/51) had a median switch time of 101 days while those who started after 2010 (28/51) had a median of 141 days. It is possible this relates to increasing concerns about PML in natalizumab-treated patients after 2010 and a desire to avoid asymptomatic patients being treated with alemtuzumab. Our study did not assess the measures taken during the switch period to exclude PML before alemtuzumab was commenced but 23 of the 36 patients (64%) with available MRI data had an MRI brain during the Switch period and no cases of PML have been reported in this cohort.

Table 3-39: Efficacy cohort characteristics by duration of Switch period

	Total (Efficacy) cohort (N=51)	Switch period < 4 months (N=26)	Switch period > 4 months (N=25)
Baseline characteristics			
% Female	40 (78%)	24 (92%)	16 (64%)
Mean Age at diagnosis (SD)	29 (8)	30 (10)	28 (6)
Median Disease duration at diagnosis¶ [Years] (IQR)	0.34 (1.1)	0.29 (1.2)	0.50 (0.7)
Mean baseline EDSS (SD)	2.3 (1.7)	2.5 (1.9)	2.2 (1.5)
Mean number of relapses pre-diagnosis (SD)	1.9 (1.3)	1.7 (1.0)	2.1 (1.6)
Patients with no co-morbidities	39 (76.5%)	17 (65%)	22 (88%)
Employed	27 (66%) (n=41)	14 (67%) (n=21)	13 (65%) (n=20)
Driving	18 (78%) (n=23)	4 (50%) (n=8)	14 (93%) (n=15)
Pre-Natalizumab			
Number of DMTs used (%) 0 1 2 3 4 5	 20 (39%) 22 (43%) 4 (8%) 3 (6%) 1 (2%) 1 (2%)	 11 (42%) 10 (38%) 3 (12%) 2 (8%) 0 0	 9 (36%) 12 (48%) 1 (4%) 1 (4%) 1 (4%) 1 (4%)
Main reason for stopping DMT(s) (%)	Efficacy (49%) Side effects (40%)	Efficacy (48%) Side effects (48%)	Efficacy (55%) Side effects (30%)
Mean duration of Pre-Natalizumab phase [years] (SD)	4.2 (4.4)	3.5 (3.8)	5.0 (5.0)
Mean ARR	0.57 (0.47,0.68 95% CIs)	0.58	0.56
Mean EDSS before Natalizumab treatment (last available) (SD)	3.4 (2.0)	3.0 (1.7)	3.9 (2.4)
MRI Brain (n=8)	(n=15)		(n=7)
Mean Gd+ lesions (SD) <u>Supratentorial lesion load (%)</u> None Minor Moderate Marked <u>Infratentorial (%)</u>	7.3 (13.5) 10 Gd scans 0 2 (13%) 2 (13%) 11 (73%)	0.2 (0.45) 5 Gd scans 0 1 (13%) 2 (25%) 5 (63%)	14.4 (16.8) 5 Gd scans 0 1 (14%) 0 6 (86%)

None	4 (27%)	2 (25%)	2 (29%)
Minor	2 (13%)	1 (13%)	1 (14%)
Moderate	5 (33%)	4 (50%)	1 (14%)
Marked	4 (27%)	1 (13%)	3 (43%)
Global atrophy (%)			
None	9 (60%)	6 (75%)	3 (43%)
Minor	6 (40%)	2 (25%)	4 (57%)
Moderate	0	0	0
Marked	0	0	0
Callosal atrophy (%)			
None	8 (53%)	6 (75%)	2 (29%)
Minor	4 (27%)	1 (13%)	3 (43%)
Moderate	2 (13%)	1 (13%)	1 (14%)
Marked	1 (0.7%)	0	1 (14%)
Natalizumab Treatment			
Mean age at 1st Natalizumab Treatment (SD)	34 (8.9)	33 (10.8)	34 (6.1)
Mean disease duration at 1st Natalizumab Treatment* [years] (SD)	4.6 (4.4)	4.5 (5.2)	4.7 (3.4)
Median duration Natalizumab Treatment [years] (IQR)	1.2 (2.1)	0.8 (1.6)	1.8 (2.0)
Median number of Natalizumab infusions (IQR, range)	12 (19, 2-91)	11 (17, 2-91)	14 (20, 4-64)
Main reasons for stopping Natalizumab	Efficacy (58%) Adverse events (19%) PML Concern (19%)	Efficacy (67%) Adverse events (22%)	Efficacy (29%) PML concern (25%) Not documented (25%)
Number of JCV +ve patients	16 (70%) (n=23)	7 (54%) (n=13)	**9 (90%)*** (n=10)
Number of patients with Natalizumab Neutralising Abs	6 (14%) (n=44)	4 (15%) (n=26)	2 (11%) (n=19)
Switch period			
Mean duration Switch period [days] (SD)	163 (139) [5.4 months]	78 (25) [2.6 months]	**273 (162)* [9.1 months]**
Patients using DMTs during Switch (%)	7 (14%)	0	**7 (28%)***
Alemtuzumab Treatment			
Mean duration of follow-up after 1st Alemtuzumab dose (years) (SD)	4.8 (2.1)	5.5 (5.1)	**4.1 (1.9)***
Mean number of Alemtuzumab courses (SD, range)	2 (1, 1-4) (6 missing)	2.3 (0.6,2-4) (1 missing)	2.2 (0.6, 1-3) (5 missing)

Number of Alemtuzumab Courses 1 2 3 4 Unknown	 2 (4%) 30 (59%) 12 (24%) 1 (2%) 6 (12%)	 0 18 (67%) 7 (26%) 1 (4%) 1 (4%)	 2 (8%) 12 (50%) 5 (21%) 0 5 (21%)
Mean age at first Alemtuzumab dose (SD)	36 (8.6)	35 (10.9)	37 (4.9)
Mean disease duration at 1st Alemtuzumab dose* (years) (SD)	6.7 (5.2)	5.8 (5.3)	8.7 (4.7)
Number of patients using DMTs after Alemtuzumab (%)	16 (31%)	11 (42%)	5 (20%)

¶ Disease duration = since first symptom (not diagnosis)

***p<0.05**

DMT use during the Switch period

Only 7 patients (14%) used DMTs during the Switch period in the efficacy cohort and all were in the longer switch duration group (>4 months). Fingolimod was used in 6 patients and glatiramer acetate in one. Additionally, one patient, having used fingolimod, briefly returned to natalizumab for 3 months before switching to alemtuzumab within 2 months of stopping. As noted above, there was a difference by length of switch and the reasons for stopping natalizumab where those with a shorter switch period had more efficacy concerns while those with longer switch had more PML concerns. However, this was due to the small group of patients with longer switch periods who had been given DMTs, all of whom had stopped natalizumab due to PML concern. Once these were removed there was no significant difference (fishers exact test p=0.12) between the longer and shorter switch groups in terms of the reason for stopping natalizumab.

Safety outcomes by Switch duration

The number of adverse events in the Alemtuzumab treatment phase differed significantly by the duration of the Switch period, with a smaller number of adverse events among those with the longer duration. There were more adverse

events in the Post-alemtuzumab phase among those with the longer switch but the difference was not significant.

Table 3-40: Adverse events after Alemtuzumab treatment by switch duration

Phase	Total adverse events	Switch under 4 months (n=26)	Switch over 4 months (n=14)	Bridging DMT (n=6)	p (two switch groups)
Alemtuzumab RX	Mean±SD	1.69±1.64	0.57±0.85	0.83±0.41	
	Median±IQR	1.0±1.25	**0±1.25**	1±0.25	p<0.05*
	Range (min,max)	0,6	0,2	0,1	
Post Alemtuzumab	Mean±SD	0.19±0.40	0.39±1.24	0	
	Median±IQR	0.0±0.0	0±0.0	0	p=0.60*
	Range (min,max)	0,1	0,5	0,0	

* Mann Whitney test

Efficacy outcomes by Switch duration

Relapses

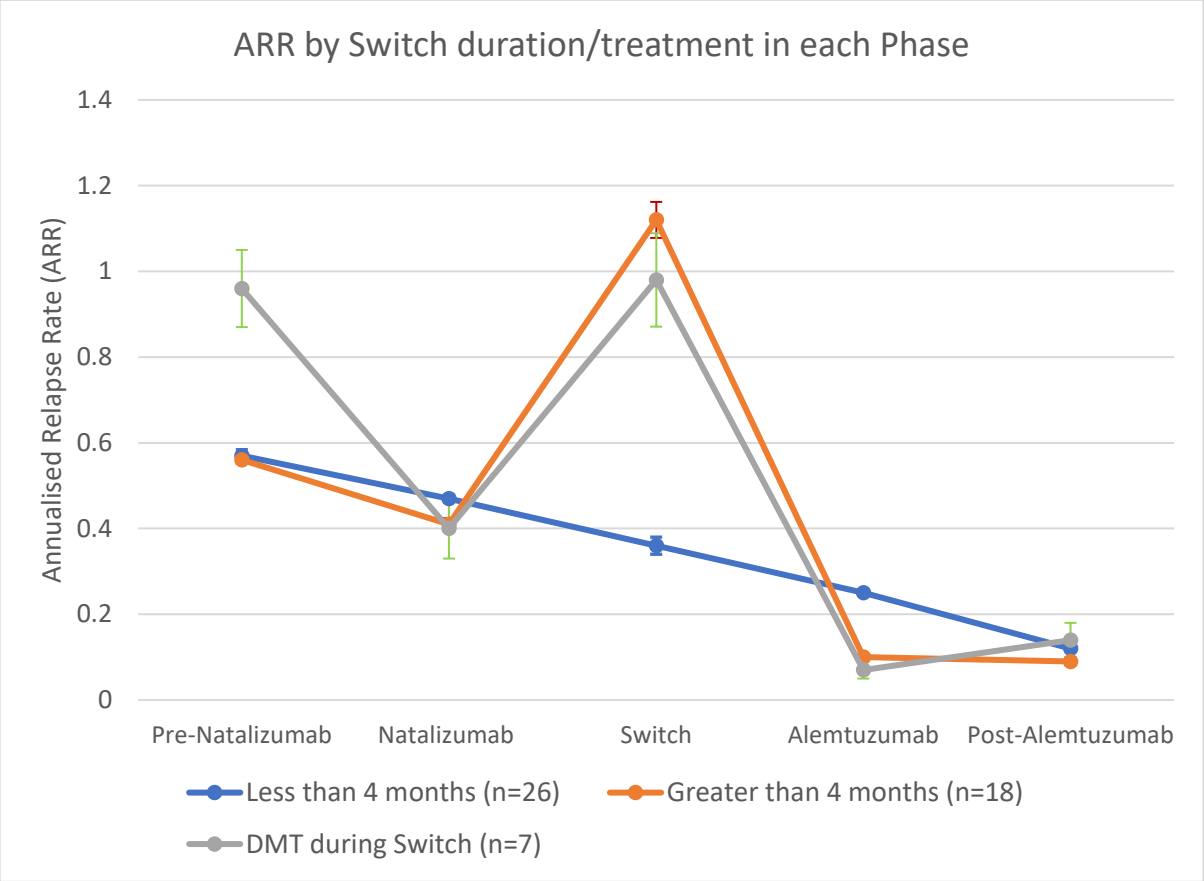
There was a significantly higher relapse rate during the Switch period in patients taking more than 4 months to switch between natalizumab and alemtuzumab, whether a 'bridging' DMT was used during this period or not, compared to patients switching in less than 4 months (See **Table 3-41** and **Figure 3-32**). Having had a lower ARR during the Natalizumab treatment phase, the longer switch groups had a significant increase during the Switch period, whilst this actually fell in the shorter switch group. Both groups had lower ARR in the Alemtuzumab treatment and Post-alemtuzumab phases, the longer switch group now being significantly lower in these phases than the shorter switch group. The difference in magnitude in the ARR is most pronounced during the Switch period, however.

Table 3-41: ARR (95% CIs) in each treatment phase by Switch duration [N=51]

Switch Duration	Pre-natalizumab	Natalizumab treatment	Switch period	Alemtuzumab treatment	Post-alemtuzumab
< 120 days (n=26)	0.57 (0.56, 0.58)	0.47 (0.46 ,0.48)	**0.36 (0.34, 0.38)**	0.25 (0.245, 0.255)	0.12 (0.11, 0.13)
> 120 days (n=18)	0.56 (0.55, 0.57)	**0.41 (0.40, 0.42)**	1.12 (1.08, 1.16)	**0.10 (0.09, 0.11)**	**0.09 (0.08, 0.10)**
DMT used during Switch [all > 120 days (n=7)]	0.96 (0.87, 1.05)	0.40 (0.33, 0.47)	0.98 (0.87, 1.09)	0.07 (0.05, 0.09)	0.14 (0.10, 0.18)

NB: Significantly lower ARR (<4 months vs >4 months) highlighted in **bold**

Figure 3-32: Annualised Relapse rate in each phase by Switch duration



The shorter switch group had a significantly longer relapse-free period after stopping natalizumab than the longer switch group but there was no significant difference in the proportion of patients remaining relapse-free until the end of follow-up between the groups. **Table 3-42** shows that a greater proportion of the shorter switch group remained relapse-free until the end of follow-up (50%) compared to the longer switch group (33%) but the difference was not statistically significant (fisher's exact test p=0.54). Overall, 22/51 (43%) of the Efficacy cohort had no further relapses after switching from natalizumab to alemtuzumab until the end of follow-up, with over 4 years of follow-up on average in both groups, suggesting a lasting effect of alemtuzumab in a significant proportion of patients irrespective of switch length.

Table 3-42: Proportion relapse-free until end of follow-up by Switch duration

Switch duration	Number (%) of patients with no relapses after natalizumab completion until end of follow-up
<120 days (n=26)	13 (50%)
>120 days (n=18)	6 (33%)
Bridging DMT used [>120 days (n=7)]	3 (43%)
Total	22 (43%)

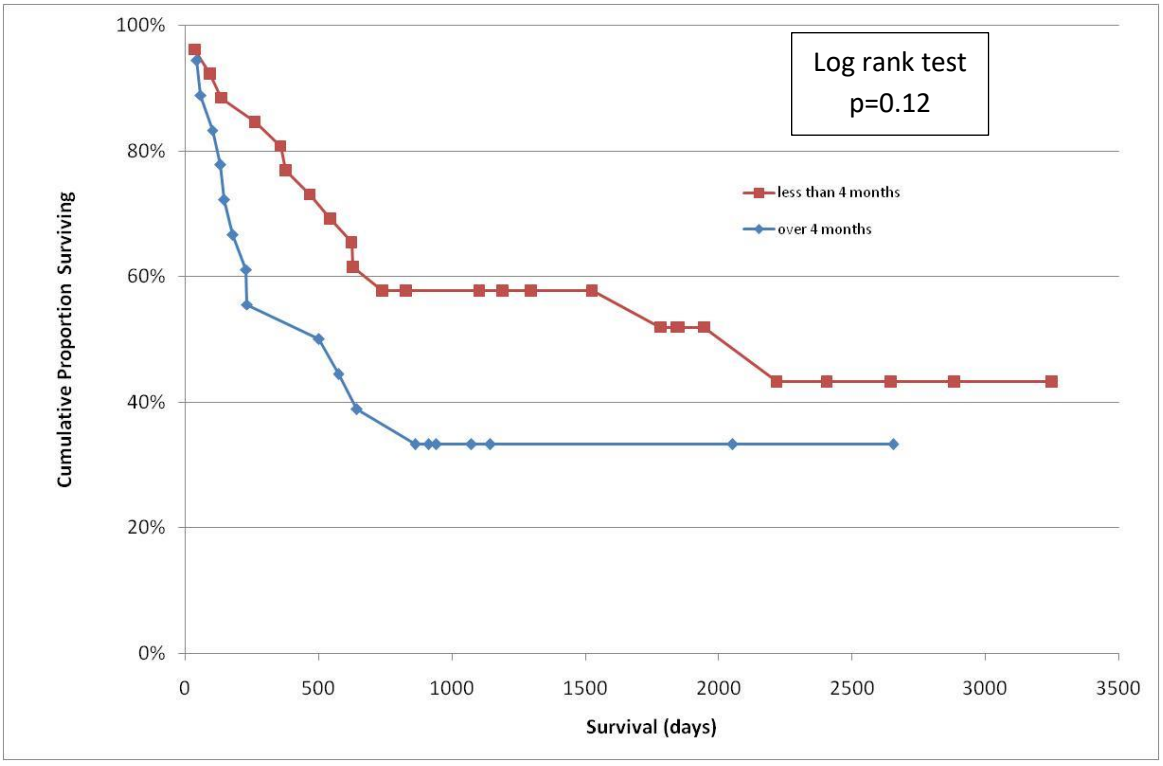
Analysis was undertaken to evaluate time to first relapse after stopping natalizumab, comparing the shorter and longer switch groups. Median survival time was 740 days overall, but was significantly longer for the <4 months group (1146 days) than the >4 months group (537 days) (Mann Whitney p=0.038) as shown in **Table 3-43**. Using a Kaplan-Meier analysis, again the shorter switch group had longer relapse-free survival as shown in **Figure 3-33**, but this was not statistically significant (log rank test p=0.120). There was also no significant difference when the DMT bridging group was included.

Table 3-43: Relapse-free survival after natalizumab by Switch duration

Survival time	Overall (n=51)	Switch <4 months (n=26)	Switch >4 months (n=18)	Bridging DMT (n=7)
Mean±SD	995±912	1318±1005	693±716	575±575
Median±IQR	740±1339	**1146±1570***	537±831	186±1100
Range (min,max)	35,3249	35,3249	42,2655	89,1306

***p<0.05**

Figure 3-33: Kaplan-Meier relapse-free survival by group switch length



Focusing on outcomes after the Alemtuzumab treatment, there was no significant difference in the percentage of patients with no relapses by switch length in either the Alemtuzumab treatment or Post-alemtuzumab phases (See **Table 3-44**). Both annualised relapse and disabling relapse rates were lower for those with longer switch length in both the Alemtuzumab and Post-alemtuzumab phases, similar to relapse rates, but none of the differences were significant. The group of patients treated with a DMT had the lowest overall annualised relapse rate in both the Alemtuzumab treatment and Post-alemtuzumab phases

but this was based on a small number of patients with a small number of events and the confidence intervals are therefore understandably wide, making this finding of uncertain significance. The longer switch group had had a higher ARR during the Switch period (See **Figure 3-32**) so the reduced relapses thereafter could be regression to the mean or that they had a greater response to alemtuzumab treatment than those switching in a shorter period.

EDSS

There was no significant difference between the two main switch groups in EDSS AUC for either the Alemtuzumab treatment or Post-alemtuzumab phases. However, disability levels were numerically higher for those with the shorter switch period and the absolute values were better for patients with a longer switch period and using a bridging DMT but this did not reach significance (p=0.06) as shown in **Table 3-45**. Also shown, there was no significant difference by switch duration in the AUC categorisation by switch duration in either the Alemtuzumab or Post-alemtuzumab phases ('improved', 'unchanged', 'worse') although those with a longer switch duration appeared to have a lower proportion with net worsening disability by this method, again this was not statistically significant.

Table 3-44: Relapse rate after Alemtuzumab treatment by switch duration

Phase	Alemtuzumab Rx			Post-Alemtuzumab		
Length of switch period	Switch under 4 months (n=26)	Switch over 4 months (n=18)	DMT used (n=7)	Switch under 4 months (n=25)	Switch over 4 months (n=18)	DMT used (n=7)
No. of patients with no relapses (%) 95%CI††	16 (62%) (43%,78%)	14 (78%) (55%,91%)	6 (86%) (49%,97%)	17 (68%) (48%,83%)	13 (72%) (49%,88%)	6 (86%) (49%,97%)
Total number of relapses (max per patient)	13 (2)	4 (1)	1	17 (6)	8 (3)	1
Total patient yrs in phase	52.0	36.0	14.0	89.1	45.3	7.3
Overall annualised relapse rate* (95% CI†)	25.0 (13.3,42.8)	11.1 (3.0,28.4)	7.1 (0.2,39.8)	19.1 (11.1,30.5)	17.7 (7.6,34.8)	13.7 (0.3,76.3)
Total number of disabling relapses (max per patient)	3 (1)	1 (1)	0	11 (5)	4 (2)	1
Overall annualised disabling relapse rate* (95% CI†)	5.8 (1.2,16.9)	2.8 (0.1,15.5)	0	12.3 (6.2,22.1)	8.8 (2.4,22.6)	13.7 (0.3,76.3)

*Per 100 patient years †Using Byar's method. ††Using the Wilson Score method

Table 3-45: Annualised Total AUC (baseline adjusted) and AUC categorisation by switch duration

		Switch under 4 months (n=24)	Switch over 4 months (n=16)	DMT used during Switch (n=7)	p (two switch groups)
Max EDSS per patient					
Pre Natalizumab	Mean±SD	3.2±1.8	4.2±2.4	4.2±2.4	
	Median±IQR	3±2.3	4.3±4	4.3±5	p=0.25 NS
	Range (min,max)	(1,7.5)	(1,8.5)	(2,8)	
Natalizumab Treatment	Mean±SD	4.2±2.1	4.9±2.3	4.0±2.6	
	Median±IQR	4.5±4	6.0±3.5	3.5±4	p=0.26 NS
	Range (min,max)	(1,7)	(1,8)	(0,7)	
Annualised Total AUC (baseline adjusted)		Switch under 4 months (n=20)	Switch over 4 months (n=12)	Fingolimod (n=6)	p (two switch groups)
Alemtuzumab Treatment	Mean±SD	-0.31±1.24	-0.32±0.85	-0.78±0.91	
	Median±IQR	0.0±1.63	-0.20±0.55	-0.78±1.50	p=0.72 NS*
	Range (min,max)	-3.90,1.37	-2.35,0.99	-2.23,0.25	
	AUC categorisation				
	Net improved (%)	7 (35%)	3 (25%)	3 (50%)	p=0.40 NS*
	Net unchanged (%)	8 (40%)	9 (67%)	3 (50%)	
	Net Worse (%)	5 (25%)	1 (8%)	0	
		(n=19)	(n=13)	(n=5)	
Post Alemtuzumab	Mean±SD	0.67±0.96	0.16±0.74	0.39±0.88	
	Median±IQR	0.35±1.36	0±0.46	0±0.98	p=0.06 NS†
	Range (min,max)	-0.73,2.55	-0.95,2.14	0,1.96	
	AUC categorisation				
	Net improved (%)	1 (5%)	2 (15%)	0	p=0.27 NS†
	Net unchanged (%)	10 (53%)	9 (69%)	4 (80%)	
	Net Worse (%)	8 (42%)	2(15%)	1 (20%)	

* Mann Whitney test †Fishers exact test.

The analyses above include outcomes after the Switch period but, as shown in **Figure 3-34**, there was no significant difference in EDSS in any treatment phase based on Switch duration. There were fewer EDSS scores available in all treatment phases in the longer switch group, however, particularly after the Switch phase, as shown in **Table 3-46**. The longer switch group had significantly shorter available follow-up [mean 4.1 vs 5.5 years (See **Table 3-39**)] which likely explains this, but the lack of available EDSS data in this group limits conclusions regarding comparisons between the groups on all measures of disability using EDSS in the study.

Figure 3-34: EDSS in each treatment phase by Switch duration (N=51)

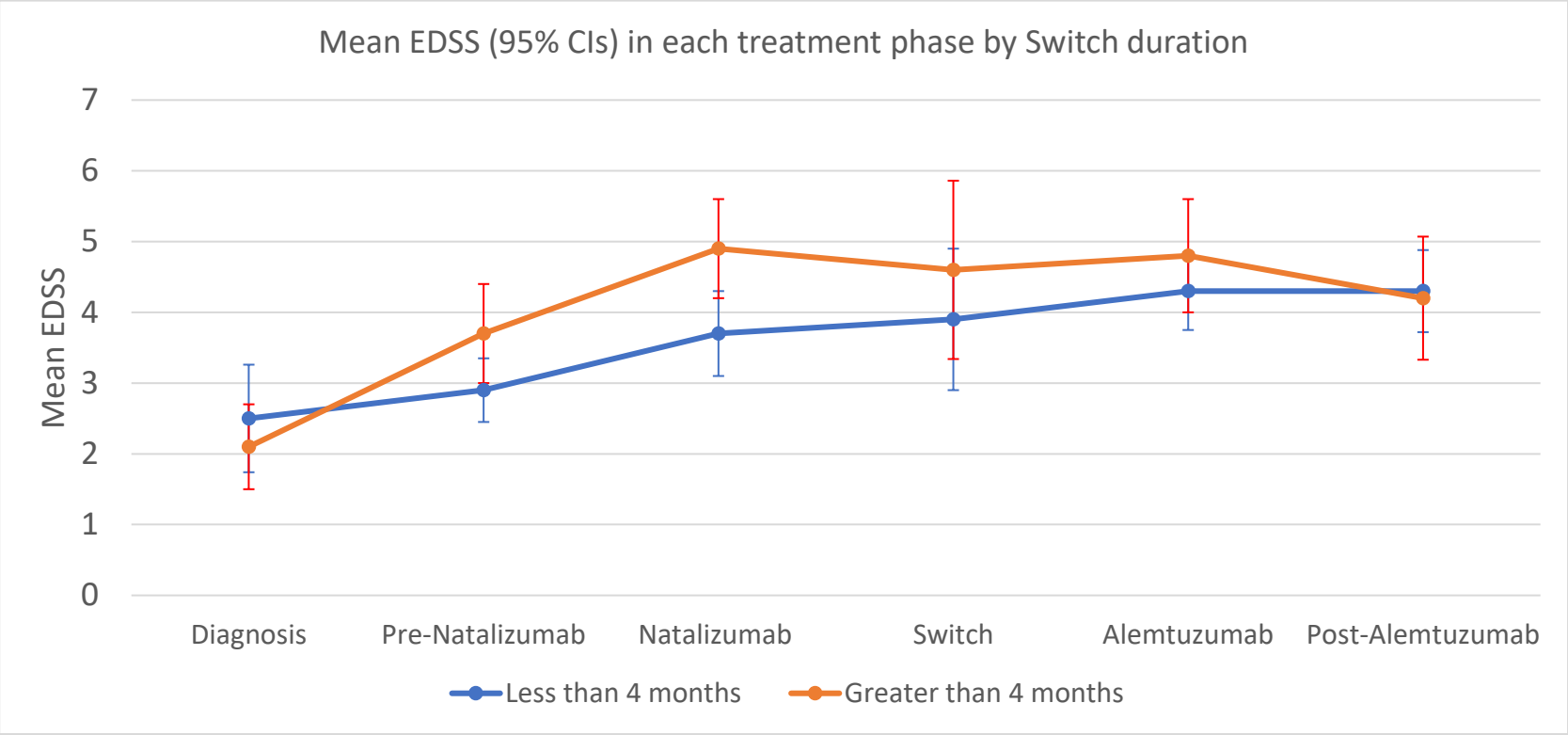


Table 3-46: Number of available EDSS scores in each treatment phase by Switch duration

	Diagnosis	Pre-natalizumab	Natalizumab treatment	Switch	Alemtuzumab treatment	Post-alemtuzumab
Switch <4 months	24	55	41	15	51	51
Switch >4 months	15	34	37	14	25	24

Combining relapse and EDSS data, 11/51 patients (22%) had no further relapses or increases in EDSS after switching from natalizumab to alemtuzumab, defined as clinical NEDA. In those where relapses or EDSS increase did occur, relapse was the more common disease activity (60% ;24/40) compared to EDSS (40%; 16/40). In terms of the effect of switch duration on maintaining clinical NEDA until the end of follow-up, again there was no significant difference in the proportions from the longer or shorter switch groups but the shorter switch group had significantly longer event-free survival.

Fewer in the over 4-month switch group remained NEDA (11%) compared to the other two groups but the difference was not significant (fisher's exact test p =0.51), as shown in **Table 3-47**.

Table 3-47: Proportion Clinical NEDA until end of follow-up by Switch duration

Switch duration	Number (%) of patients with Clinical NEDA after natalizumab completion until end of follow-up
<120 days (n=26)	7 (27%)
>120 days (n=18)	2 (11%)
Bridging DMT used [>120 days (n=7)]	2 (29%)
Total	11 (22%)

Median survival time with Clinical NEDA after natalizumab discontinuation was 467 days overall but was significantly longer for the <4 months group (650 days) in comparison to the >4 months group (203 days) (Mann Whitney p=0.014) as

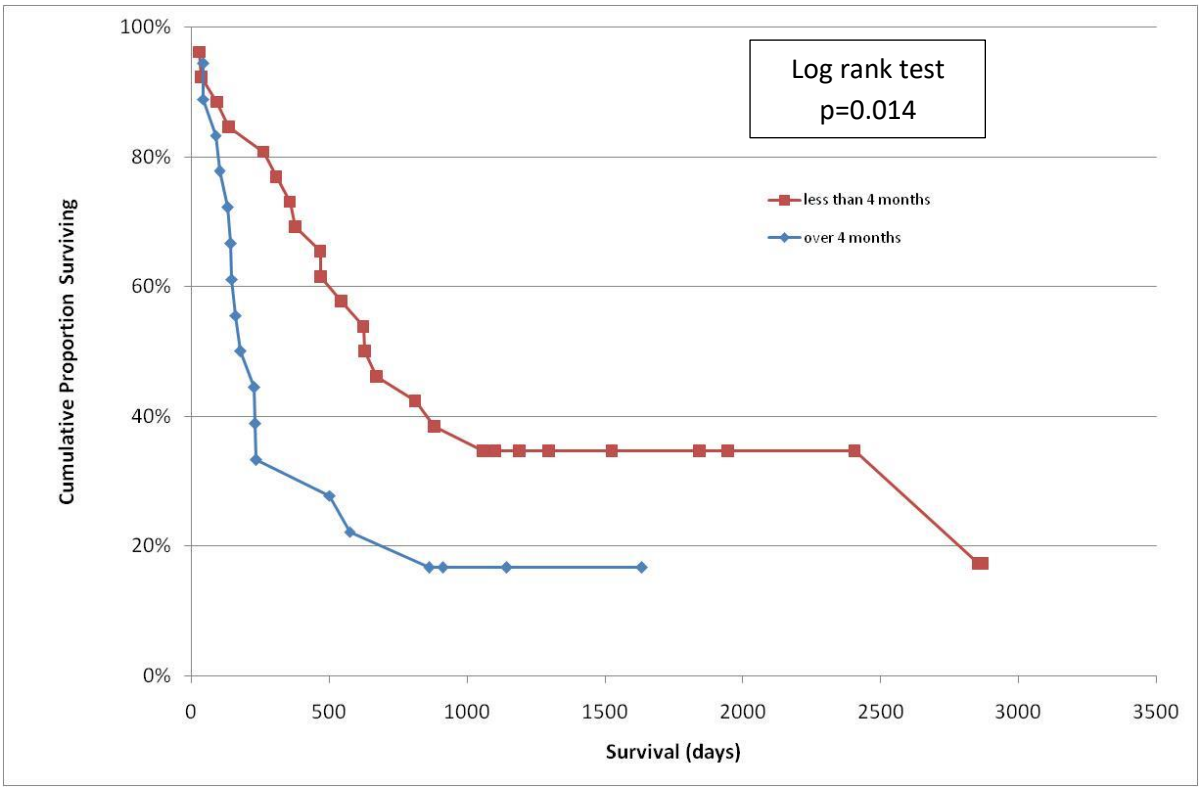
shown in **Table 3-48**. There was also a statistically significant difference using Kaplan-Meier analysis, illustrated in **Figure 3-35**, with longer maintenance of Clinical NEDA in the shorter switch group. The trend remained the same when the bridging DMT group was included but the difference was no longer statistically significant (log rank test p = 0.054).

Table 3-48: Clinical NEDA survival by switch duration (N=51)

Survival time	Overall (n=51)	Switch <4 months (n=26)	Switch >4 months (n=18)	Bridging DMT (n=7)
Mean±SD	693±724	953±833	408±448	455±546
Median±IQR	467±961	**650±1009***	203±522	186±1100
Range (min,max)	29,2871	29,2871	42,1633	89,1306

***p<0.05 (vs. Switch > 4months group)**

Fig 3-35: Kaplan-Meier analysis of Clinical NEDA survival by switch length



No Evidence of Disease Activity (NEDA)

Further Kaplan Meier analysis was conducted to investigate whether the length of switch period between natalizumab and alemtuzumab had any effect on the amount of time the patient remained free of all measures of disease activity available (NEDA), defined as no relapses, new MRI lesions or increase in EDSS. Survival was from the date of the last natalizumab infusion to an end date that was either the first instance of disease activity or the end data of the study for censored cases.

There were 51 patients in the analysis of which 10 (20%) remained NEDA at the end of the study period. The most common disease activity event was relapse [59% (24/41)] followed by EDSS increase [29% (12/41)] with only 12% due to new MRI lesion (5/41).

As shown in **Table 3-49**, fewer in the longer switch duration group remained NEDA (11%) compared to the other two groups but the difference was not significant (fisher's exact test p =0.55).

Table 3-49: Proportion NEDA until end of follow-up by Switch duration

Switch duration	Number (%) of patients with NEDA after natalizumab completion until end of follow-up
<120 days (n=26)	6 (23%)
>120 days (n=18)	2 (11%)
Bridging DMT used [>120 days (n=7)]	2 (29%)
Total	10 (20%)

Median survival time was 331 days overall but the <4 months switch group had significantly higher median survival than the >4 months switch group (Mann Whitney p=0.025) as shown in **Table 3-50**.

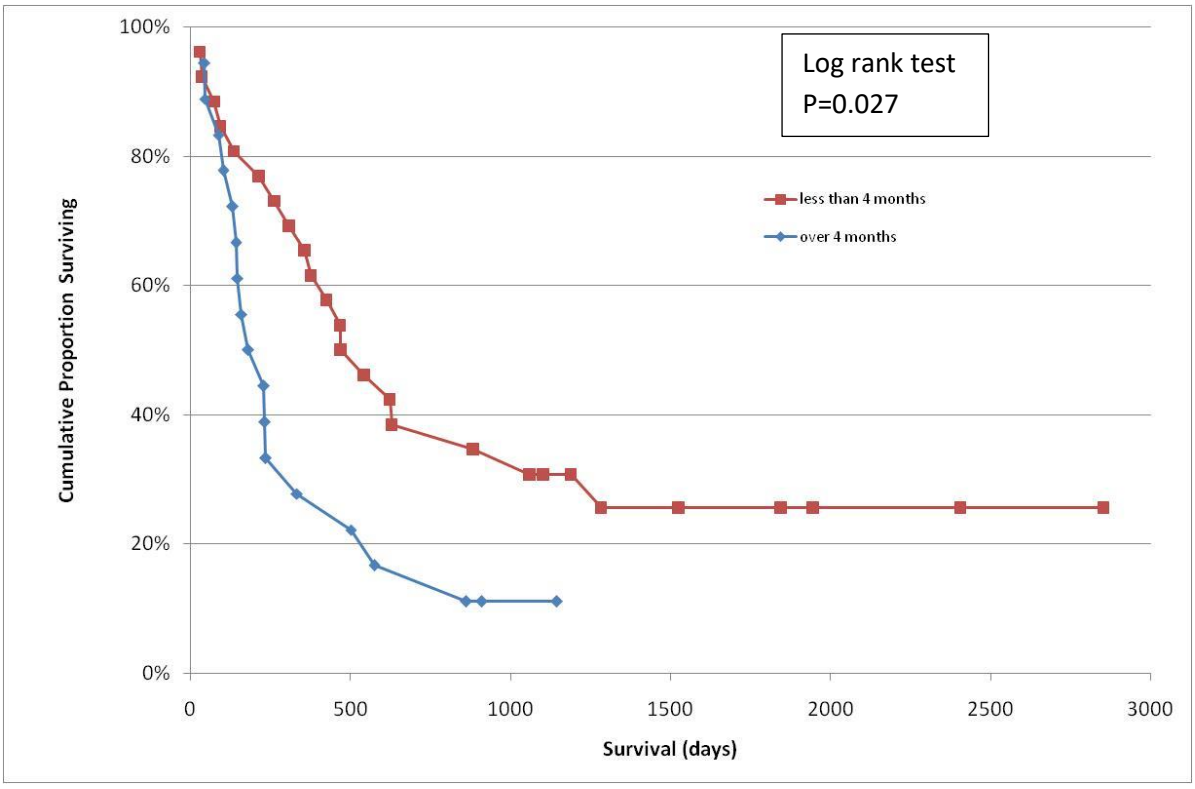
Table 3-50: NEDA survival by Switch duration

Survival time	Overall (n=51)	Switch <4 mths (n=26)	Switch >4 mths (n=18)	Bridging DMT (n=7)
Mean±SD	595±646	813±766	336±329	455±546
Median±IQR	331±775	**506±962***	203±396	186±1100
Range (min,max)	29,2854	29,2854	42,1144	89,1306

***p<0.05 (vs. >4-month Switch group)**

The Kaplan-Meier survival curve shown in **Figure 3-36** illustrates that the <4-mth switch group had significantly longer NEDA survival times than the >4mth group (log rank test p=0.027). Including the Bridging DMT group, the shorter switch group still had longer survival but the difference was not statistically significant.

Figure 3-36: Kaplan-Meier analysis of NEDA survival by switch length



Symptomatic therapy by Switch duration

After the Switch phase, there were differences in the number of symptomatic treatments in the Alemtuzumab and Post-alemtuzumab treatment phases by

duration of switch. The mean number of treatments used was lower for those with the longer switch duration: this difference was significant for the Alemtuzumab treatment phase but not for the Post-alemtuzumab phase (See **Table 3-51**).

Table 3-51: Symptomatic therapy by switch duration

Phase	Total Symptomatic therapies	Switch under 4 months (n=27)	Switch over 4 months (n=24)	sig
Alemtuzumab Treatment	Mean±SD	1.74±1.77	0.75±1.23	p<0.05*
	Median±IQR	1.0±3.0	0±1.75	
	Range (min,max)	0,7	0,4	
Post Alemtuzumab	Mean±SD	1.48±2.00	0.79±1.62	p=0.08 NS*
	Median±IQR	1.0±2.0	0±1.0	
	Range (min,max)	0,8	0,6	

* Mann Whitney test

Switch due to Natalizumab efficacy failure only (N=25)

In the efficacy cohort (N=51), patients were switched from natalizumab to alemtuzumab mainly due to safety concerns (PML risk) or lack of efficacy. The longer switch group (when those treated with a Bridging DMT during the Switch period are included) were more likely to stop natalizumab due to PML concerns rather than efficacy and appeared to have a less active disease phenotype on some measures, hence long-term outcomes may reflect selection bias rather than the effects of the switching process. In order to address this, the cohort of patients switching from natalizumab to alemtuzumab because of lack of efficacy with natalizumab are evaluated separately here.

Twenty-five patients switched from natalizumab to alemtuzumab due to lack of efficacy. **Table 3-52** outlines this cohort and compares those who switched in less than or more than approximately 3 months (91 days) and compares their subsequent outcomes. The 3-month cut-off was chosen to provide roughly equal numbers in each group.

Table 3-52: **Natalizumab efficacy failure cohort and outcomes by Switch duration**

	All (N=25)	Switch period < 3 months (n=11)	Switch period > 3 months (n=14)
Baseline characteristics			
% Female	20 (80%)	10 (91%)	10 (71%)
Mean Age at diagnosis (SD)	30 (7.7)	31 (7.6)	29 (7.8)
Median Disease duration at diagnosis* [Years] (IQR)	0.41 (1.2)	0.77 (1.2)	0.39 (0.54)
Mean baseline EDSS (SD)	2.7 (1.9)	3.1 (2.0)	2.3 (1.8)
Mean number of relapses pre-diagnosis (SD)	2.0 (1.0)	1.9 (1.1)	2.0 (1.0)
Patients with no co-morbidities	17 (68%)	7 (64%)	10 (71%)
Pre-Natalizumab			
Number of DMTs used (%) 0 1 2 3	11 (44%) 12 (48%) 1 (4%) 1 (4%)	3 (27%) 6 (55%) 1 (9%) 1 (9%)	8 (57%) 6 (43%) 0 0
Main reason for stopping DMT(s) (%)	Efficacy (59%) Side effects (29%)	Efficacy (55%) Side effects (36%)	Efficacy (67%) Side effects (17%)
Mean duration of Pre-Natalizumab phase [years] (SD)	4.2 (5.2)	4.9 (4.9)	3.7 (5.6)
Mean ARR	0.51	0.71	0.29
Mean EDSS before natalizumab treatment (last available) (SD)	3.3 (2.0)	3.4 (1.9)	3.2 (2.2)
Natalizumab Treatment			
Mean age at 1st Natalizumab Treatment (SD)	33.9 (8.5)	35.9 (8.4)	32.4 (8.5)
Median duration Natalizumab Treatment [years] (IQR)	0.97 (1.79)	1.1 (2.3)	0.84 (1.2)
Median number of Natalizumab infusions (IQR, range)	12 (15)	20 (10.3)	10.5 (6.8)
Number (%) of JCV +ve patients	6 (55%) [n=11]	5 (71%) [n=7]	1 (25%) [n=4]
Switch period			

Mean duration Switch period [days] (SD)	117 (93.1) [3.9 months]	67 (15.4) [2.2 months]	155 (110) [5.2 months]
Patients using DMTs during Switch (%)	0	0	0
Alemtuzumab Treatment			
Mean duration of follow-up after 1st Alemtuzumab dose (years) (SD)	5.4 (2.0)	4.8 (2.0)	5.8 (2.0)
Mean number of Alemtuzumab courses (SD, range)	2.4 (0.58)	2.4 (0.50)	2.5 (0.66)
Number of Alemtuzumab Courses 2 3 4 Unknown	 15 (60%) 8 (32%) 1 (4%) 1 (4%)	 7 (64%) 4 (36%) 0 0	 8 (57%) 4 (29%) 1 (7%) 1 (7%)
Mean age at first alemtuzumab dose (SD)	36.1 (8.6)	38.1 (8.7)	33.8 (7.9)
Number of patients using DMTs after alemtuzumab (%)	11 (44%)	6 (55%)	5 (36%)
Post-alemtuzumab Outcomes			
Safety			
Number of patients with any AE (%)	Alem: 18 (72%) Post-Alem: 5 (20%)	Alem: 8 (73%) Post-Alem: 2 (18%)	Alem: 10 (71%) Post-Alem: 3 (21%)
Number of patients with autoimmune thyroid disease (%)	5 (20%)	2 (18%)	3 (21%)
Number of patients with infection (%)	5 (20%)	3 (27%)	2 (14%)
Efficacy			
Alemtuzumab Treatment mean ARR	0.26	0.27	0.21
Post-alemtuzumab mean ARR	0.13	0.08	0.17
Alemtuzumab mean EDSS (SD)	4.4 (2.0)	5.3 (1.7)	3.6 (2.0)
Post-alemtuzumab mean EDSS (SD)	4.1 (2.1)	4.9 (1.6)	3.8 (2.2)

*Disease duration = since first symptom (not diagnosis)

Again, overall, the shorter duration switch group appears to have more active disease which likely biases longer-term outcomes and may explain the clinical reasoning for the shorter switch period. The shorter switch group had

proportionally more females, with older age, longer disease duration and higher EDSS at diagnosis than the longer switch group. The longer switch group were more likely to be treated (younger) with natalizumab as a first-line therapy, yet had lower Pre-natalizumab relapse rate and EDSS in comparison to the shorter switch group. The shorter switch group had a longer period of natalizumab treatment on average, a greater proportion were JCV antibody positive and were older at the time of alemtuzumab treatment, but follow-up after alemtuzumab was shorter. A greater proportion of the shorter switch cohort required DMTs after alemtuzumab treatment in comparison to the longer switch group.

In terms of outcomes between the groups with longer or shorter switch duration who failed natalizumab because of lack of efficacy, there was little difference in safety but some (non-statistically-significant) differences in efficacy. The proportion of patients having any adverse event or secondary autoimmunity was comparable in each group although infections were proportionally higher in the shorter switch group. The shorter switch group had a lower relapse rate in the Post-alemtuzumab phase but higher mean EDSS in comparison to the longer switch group.

The numbers here are too small to draw conclusions, but this analysis likely confirms that of the larger cohort analysis, namely that clinicians switched those with more disease activity quicker to maintain ongoing treatment effect. This, in turn perhaps explains the generally poorer outcomes in the shorter switch period, rather than being able to relate this directly to the duration of the switch although this is an alternative explanation.

The Switch period: Conclusions

Our analysis focused on the duration of the switch period between natalizumab and alemtuzumab and its effect on safety and efficacy outcomes. Only 7 of the 51 patients used a Bridging DMT during the Switch period, making conclusions about the effectiveness of this approach limited. We therefore mainly compared patients by duration of the switch period, separated into less than or greater than 4 months between natalizumab stopping and alemtuzumab starting. Overall, the shorter switch group had more aggressive disease before the switch, which may explain some of the poorer outcomes during the Alemtuzumab and

Post-alemtuzumab phases, but generally the shorter switch group had longer disease-free survival on all measures of activity.

Table 3-53 summarises the differences in outcomes between the shorter and longer switch groups, with the statistically significant differences highlighted in bold. Efficacy outcomes favour the shorter switch group although there were fewer AEs and symptomatic treatments used during alemtuzumab treatment in the longer switch group. There was no significant difference in mean EDSS scores between groups in any treatment phase, however. The EDSS AUC analyses generally favoured the longer switch, including the Bridging DMT group, in the Alemtuzumab and Post-alemtuzumab phases but this was not statistically significant. When we analysed only patients stopping natalizumab because of efficacy failure (N=25), the shorter switch group (<3 months) had lower Post-alemtuzumab relapse rates but higher mean EDSS scores. The worse disability outcomes in the shorter switch groups may reflect the more active disease status preceding the switch rather than the characteristics of the switch itself, however. There were proportionally more infections during the Alemtuzumab treatment phase with the shorter switch group also, but otherwise safety outcomes were comparable.

Table 3-53: Summary of outcomes between <4months and >4 months switch duration groups (statistically significant differences highlighted in **bold**)

Favours < 4months switch	Favours > 4months switch
Lower ARR during Switch period (0.36 vs 1.12) [See Table 3-41]	**Fewer AEs during the Alemtuzumab treatment phase** (p<0.05) but no difference in Post-alemtuzumab phase [See Table 3-40]
Higher proportion relapse-free after natalizumab until end of follow-up (50% vs 33%) [See Table 3-42]	**Lower ARR in Alemtuzumab treatment phase (0.1 vs 0.25) and Post-alemtuzumab phase (0.09 vs 0.12)** [See Table 3-41]
Longer median duration to any relapse (1146 days vs 537 days) [See Table 3-43]	Median EDSS AUC lower for longer switch group in Alemtuzumab treatment phase (0 vs -0.2) and Post-alemtuzumab phase (0.35 vs 0) [See Table 3-45]
Proportion 'net improved' on EDSS AUC categorisation higher in shorter switch group (35%) vs longer switch group (25%) after Alemtuzumab treatment [See Table 3-45]	Lower proportion 'net worse' using EDSS AUC categories during Alemtuzumab (8% vs 25%) and Post-alemtuzumab (15% vs 42%) treatment phases [See Table 3-45]
Longer duration of clinical NEDA survival after switch: 650 days vs 203 days in longer switch group (p=0.014) [See Table 3-48 and Figure 3-35]	**Lower number of symptomatic treatments used in the Alemtuzumab treatment phase (0.75 vs 1.71)** but no significant difference in the Post-alemtuzumab treatment phase [See Table 3-51]
Longer survival of NEDA after natalizumab stopped – median 506 days vs 203 days [See Table 3-50 and Figure 336]	

Discussion

This retrospective observational study evaluated the safety and efficacy of switching to alemtuzumab after failure of natalizumab in patients with RRMS, using routinely collected clinical data. We identified 79 patients from 13 MS specialist centres from the UK & Ireland who had made this switch between March 2009 and January 2017. This cohort was used for safety analysis and had follow-up data for 3.6 years after starting alemtuzumab on average. A subgroup of patients (N=51) had at least 2 years' follow-up data available after starting alemtuzumab and were evaluated additionally for efficacy outcomes. In this subgroup, the switch occurred between March 2009 and October 2015 and a mean of 4.8 years of follow-up data were available from starting alemtuzumab

and 11.2 years from diagnosis. MRI data were available for 36 of the 79 patients included overall and 25 were within the efficacy subgroup and so had longer follow-up data available (mean 5.3 years from first alemtuzumab course). There were some differences in the baseline characteristics of the subgroups in comparison to the total cohort, but none of these reached statistical significance. We assessed patients in 5 retrospectively-defined phases after diagnosis and baseline features were characterised; Pre-natalizumab treatment; Natalizumab treatment; Switch period; Alemtuzumab treatment and Post-alemtuzumab treatment. The Alemtuzumab treatment period was defined for the purposes of this study as the first 2 years after alemtuzumab was started. All other phases were of variable duration and occurred at different time points. For example, the earliest diagnosis was made in 1988 and the most recent in 2016, with natalizumab used between 2005 and 2016 and alemtuzumab between 2009 and 2017 in the cohort overall. The majority of patients (57%) had been treated with DMTs prior to natalizumab but natalizumab was used first-line in the remainder, reflecting the highly active nature of the cohort. Almost half of patients (49%) stopped natalizumab because of breakthrough disease on treatment and 15% due to PML concerns because of positive JCV antibodies.

The results suggest that that switching to alemtuzumab after failure of natalizumab in patients with RRMS is safe and improves inflammatory disease control over subsequent years, reducing the need for further DMTs in most, but initial benefits on disability are not maintained in the longer term. Overall though, despite sequential use of these high-efficacy DMTs, disability worsened on average by 0.25 EDSS points per year in this cohort, though there was some discrepancy in the overall disability outcomes depending on the analysis used. Using patients with longitudinal EDSS scores, 20/35 patients (57%) had either minimal change or sustained improvement with SAD/SRD categorisation, whereas 25/35 (71%) had net worsening disability using the AUC method over the entire study period. When the parameters of the AUC evaluation were altered, there was closer agreement between these two methods but still the majority had worsening disability (21/35, 60%) using the AUC method, which is likely to be the more robust option in this study where EDSS values were not systematically available at comparable time points for all patients. The largely worsening disability occurred despite short-term improvement on all disease activity

measures during the 2-year Alemtuzumab treatment phase. However, given the selected population of patients with often treatment-refractory RRMS included, the rate of disability worsening and development of progressive disease was less than might be expected. The lack of a control group means conclusions cannot be definitive but using a pre-defined SAD definition, only 5 (14.3%) of the 35 patients with necessary data available were classified as secondary progressive after a mean duration of almost 12 years from diagnosis. This is lower than would be expected from natural history studies[12] and is lower than the 20-40% conversion rate seen with 15 and 16 year follow-up studies respectively with ß-IFN use[108,240].

Safety outcomes after switching to alemtuzumab from natalizumab were generally comparable to, or lower than, those described with its use first-line. Infections occurred in 13 patients (16%), 3 of which were considered serious (4%), and there was a single case of opportunistic CMV infection. There were no cases of PML. Autoimmune thyroid disease occurred in 14 patients (17%) but there was no other secondary autoimmunity identified. Autoimmune events occurred most frequently within the first 2 years after alemtuzumab whilst infections occurred later, perhaps counterintuitively. Malignant lymphoma occurred in 1 patient after HSCT which post-dated alemtuzumab treatment and was successfully managed. A single death from urosepsis occurred 3 years after alemtuzumab treatment in an already significantly disabled patient and was not thought related to DMT use by the patient's treating physician. Notably, there were no documented cases of stroke or arterial dissection, which is reassuring in light of recent concerns, and the safety profile overall provides confidence in keeping with the recent EMA licensing decision to re-introduce alemtuzumab after a period of re-appraisal of its safety[241].

The 2-year Alemtuzumab treatment phase was associated with the lowest levels of disease activity on most measures. ARR, SAD, SRD, EDSS AUC, net new/worsening MRI Brain lesions, global and cervical cord atrophy and NEDA-3 all favoured the Alemtuzumab treatment phase in comparison to all other treatment phases. Median EDSS was lowest in the Pre-natalizumab phase and generally increased with time despite sequential therapy although there was suggestion of plateau after alemtuzumab treatment. Similarly, the rate of new contrast-enhancing lesions was lowest before natalizumab treatment and

increased with time despite treatments. The reduced measures of disease activity during the 2-year Alemtuzumab treatment phase were maintained throughout follow-up for relapses and new MRI brain and cervical cord lesions but the others were not, including disability outcomes. The disconnect between worsening disability despite improved inflammatory disease control is well recognised in the literature of DMTs in general and its occurrence here, again, suggests that even highly-effective therapies may not prevent this. Overall, however, only 16 patients (31%) required further DMTs after treatment with alemtuzumab.

The highest levels of disease activity on almost all measures occurred during the Switch period. To evaluate the relationship of switch duration with safety and efficacy outcomes, the cohort was separated into those switching between natalizumab and alemtuzumab in more or less than 4 months (120 days). There were some Pre-alemtuzumab differences in the cohorts of patients switching between treatments in less than or greater than 4 months suggesting that the shorter switch group had more active disease, but none of these were statistically significant. Bridging DMTs were used during the Switch period only in 7 patients in the longer switch group but these were excluded from efficacy analysis in order to evaluate the effects of duration only. There was no statistically significant difference in the length of the switch based on whether natalizumab was stopped for efficacy or safety reasons but there appeared to be non-patient factors involved in decision-making with a trend to increasing switch duration after 2010, on average, but this was also not statistically significant.

Switching from natalizumab to alemtuzumab in less than 4 months was associated with greater freedom from inflammatory disease in comparison to a longer switch in this cohort. Whilst there were fewer AEs and use of symptomatic treatments during the Alemtuzumab treatment phase on average for those in the longer switch group, significantly lower relapse rates and disease-free survival were seen in the shorter switch group. Lower ARRs in the Alemtuzumab and Post-alemtuzumab treatment phases were seen in the longer switch group, however, and disability outcomes also favoured this group but were not statistically significant and could reflect better prognostic factors which may actually explain the length of the switch. It appears that the shorter switch in the comparison group may have occurred due to high disease activity

preceding the switch and, hence, an attempt to avoid delay in starting effective treatment.

Study Limitations

There is no control group in our study and interventions were non-standardised over the study period. The results presented simply reflect the choices of the patients and treating clinicians at the time in each patient's disease course. The frequency and extent of clinical, safety and MRI monitoring of each patient was variable, as was the duration of each phase other than the Alemtuzumab treatment phase (which was set at 2 years to reflect its usual treatment regime and initial immunological effects) but this too is relatively arbitrary at the individual patient-level. Given the lack of a control group, conclusions regarding the outcomes in comparison to another (or no) intervention or treatment strategy cannot be drawn. Additionally, patient-level data will have been driven by disease activity: that is to say, those with more disease activity are more likely to be reviewed in clinic, have an EDSS score documented and MRI undertaken than those without. The duration of each phase will also correlate with the number of documented events. For example, the number of EDSS scores documented in the Post-alemtuzumab phase will relate to the duration of follow-up after alemtuzumab: those with shorter follow-up are less likely to have a documented EDSS/relapse and those without a relapse are less likely to have a documented EDSS score. This is indicative of the necessarily reactive nature of clinical healthcare, but may have skewed outcome measures to reflect those with more active disease. Indeed, this cohort by definition is somewhat unusual in that they are relatively treatment-refractory in order to require sequential use of highly active DMTs and therefore do not reflect a typical RRMS cohort but should be comparable to cohorts where switching from natalizumab to alemtuzumab in real-world practice is considered.

A particular issue for this study is incomplete data entry. This could be due to the lack of availability of (historical) medical records, incomplete documentation in source documents or incomplete entries into the online database used by each centre. Indeed, the online database used for data entry in this study had its own limitations. Where an event was not entered, it was

assumed it had not occurred. For example, if no relapse data were entered during a phase this was assumed to mean no relapses occurred rather than simply incomplete data entry. That said, all centres were provided with detailed instructions for data entry, bespoke to their centre, and understood the rationale and aims of the study. Additionally, centre leads and/or those entering data were asked to confirm when data entry was completed, hence this assumption seems reasonable. The lack of any documented infusion reactions with the use of alemtuzumab in the Post-alemtuzumab phase, for example, does suggest incomplete data entry however, given that the majority of patients have such reactions in all published studies.

Whilst the number of patients included is relatively small, this is the largest known longitudinal dataset of patients switching from natalizumab to alemtuzumab. The numbers are not adequate to power robust efficacy conclusions, even if this was possible in an observational cohort, but these data are the largest single-cohort evidence base from which clinicians can base decisions in this subgroup of patients with RRMS. The geographically-limited patient inclusion may restrict application to comparable populations but the number of MS centres involved makes variable local practices less likely to explain outcomes. That said, the majority of patients were from two centres (Charing Cross and Dublin) and it is possible that centre-specific populations and practices were influential in the overall results in these cases. We did not, for example, compare local treatment policies including PML exclusion before switching and the use of Listeria prophylaxis. As well as location, the era in which some of these patients were treated may limit applicability to a contemporary cohort. There have been four MS diagnostic criteria updates since the earliest diagnosed patients in this cohort were diagnosed. Additionally, approaches to treatment and, specifically, the use of alemtuzumab in clinical practice has changed. Its use as a 3rd-, 4th-, or 5th-line therapy is less likely in current UK practice, with a move to alemtuzumab being first-line in those with rapidly evolving disease rather than a late rescue therapy where the benefits are less likely to outweigh risks. It could be argued that the safety results from this study may be applicable to sequential use of alemtuzumab after other highly effective DMTs, but natalizumab's mechanism of action as modulator of immune cell trafficking, rather than immune reconstitution or induction effects, would

not support this. Application is more appropriate to the use of other immune reconstitution treatments after natalizumab or other cell-trafficking DMTs perhaps, but this is entirely speculative and requires separate investigation.

Study Strengths

This study provides observational evidence on the safety and efficacy of switching to alemtuzumab after natalizumab failure in RRMS from the largest cohort of patients with the longest follow-up available in the medical literature. This is a unique cohort due to the use of alemtuzumab prior to its licensing in the UK and Ireland. The data provide real-world evidence for a common clinical scenario which is unlikely to be investigated by a randomised controlled trial. The collaborative effort of 13 specialist centres across 4 countries is testament to the importance of the issue. We have collected a highly detailed dataset and evaluated safety and efficacy using a number of outcomes and analyses. The AUC method for EDSS evaluation over time, for example, provides the most robust and inclusive analysis possible with the available data.

Other studies assessing switching natalizumab to alemtuzumab

Our study specifically evaluates the safety and efficacy of switching from natalizumab to alemtuzumab due to lack of efficacy, tolerability or safety. There is very little literature available to provide an evidence-base for this switch and no longitudinal data in a significant dataset to inform its safety or efficacy.

The largest reported cohort of patients switching from natalizumab to alemtuzumab comes from a single centre in the USA and included 200 patients with follow-up over 6 months[242]. The outcomes are not reported in the abstract but have been presented and are taken here from a recent review[243]. No patients experienced a relapse in the first 6 months post-alemtuzumab and 43% of patients with EDSS measurements (n=162) showed improvement; 1% had EDSS worsening. MRI data was available for 160 patients and only 3 (2%) demonstrated new lesions in this time. Adverse events were considered mild and manageable overall but one death occurred in the presence of urinary infection and

unrecognised non-convulsive status epilepticus. No PML occurred. In this study, the average washout period between natalizumab and alemtuzumab was 9 weeks.

McGuigan et al. performed a retrospective audit of their patients in Dublin who had received alemtuzumab after natalizumab[244]. This included 11 patients who had been treated between 2009 and 2011. The reason for switching to alemtuzumab were either allergic reaction to natalizumab (n=3) or lack of efficacy (n=8). All patients predictably had lymphopaenia within the first month after alemtuzumab treatment. Reduced white cell count persisted for 3 months (without neutropaenia) in 4 patients but this had resolved by month 5 post-alemtuzumab. In terms of safety, no serious infections occurred and 1 patient developed hypothyroidism (9 months post-treatment). ARR was reduced from a mean of 2.5 on natalizumab to 0.1 in the first year after alemtuzumab treatment. There were no new contrast-enhancing or T2 lesions identified within the first year of treatment, despite activity whilst on natalizumab (8 had increased T2 lesions and 4 had contrast-enhancing lesions).

A similar evaluation of patients switching from natalizumab to alemtuzumab has been undertaken in Canada[245]. This multicentre retrospective case record-review included 13 patients treated with this sequence in Canada since alemtuzumab was approved for use there in December 2013. All patients switched to alemtuzumab because of inadequate disease control on natalizumab. The definition of this treatment failure is not reported. Mean disease duration was 9.2 years (range 1.5-28) and mean natalizumab treatment duration was 35 months (7-132) before the switch. On average, there was a gap of 13.2 weeks (4-72) between therapies as a washout period and patients had follow-up data available for 1 year (range 4-18 months). EDSS was reduced from 3.9 (0-6) on natalizumab to 3.6 (0-6) after alemtuzumab. Similarly, ARR reduced from 1.8 to 0.2. There was one case of hypothyroidism, one case of cystitis and one oral candida infection. There were no unexpected adverse events and no serious infections but almost all patients reported infusion-related reactions. Disability progression occurred in one patient within the first year after alemtuzumab treatment but the authors note their 22-year disease duration and that this may the missed treatment ‘window of opportunity’ in this case.

The most recent study evaluating switching from natalizumab to alemtuzumab was published after we completed data collection and was single-centre study from Italy[246]. This included only patients switching from natalizumab because of PML concern (N=16) and follow-up was less than 6 months for most, but efficacy and safety outcomes were also positive. No new MRI lesions, relapses or worsening disability occurred during this short follow-up in those with data available. In CARE-MS2, approximately 3% of patients in both the alemtuzumab and INF-β1agroups had previously had natalizumab (n=20) but the reason for its discontinuation is not stated in the publication and discontinuation at least 6 months before randomisation was an inclusion criterion[120]. Notably, a prospective study on switching from natalizumab to alemtuzumab is underway in the United States presently with the aim of completion at the end of 2019 (NCT03135249).

There were no cases of PML in our study, but a single fatal case has been reported in a patient who switched from natalizumab to alemtuzumab where PML MRI lesions were retrospectively identified and, hence, the condition predated alemtuzumab use and was attributed to natalizumab[243]. There have been no other cases of PML using alemtuzumab reported otherwise for RRMS.

The appropriate 'washout' period between stopping natalizumab and starting alemtuzumab remains controversial. In CARE-MS II, the 3% of patients switching from natalizumab to alemtuzumab had a 6-month washout period. Most now agree this is likely too long given the risk of disease reactivation on natalizumab discontinuation[247] and some authors advise no washout period as the risk of PML is so much lower than the risk of a severe relapse[243]. The European Medicines Agency (EMA) suggests a 12-week washout between stopping natalizumab and starting another DMT, based on its pharmacodynamic effects. A recent observational single-centre study from Italy reported a median of 70 days (range 41-99) between natalizumab and alemtuzumab in their cohort switching between the two[246]. Overall, expert opinion suggests the washout period should be no more than 1-2 months but a robust evidence-base for this is lacking. Additionally, however, it is suggested that subclinical PML infection is excluded in JCV-positive patients (using MRI and CSF for JCV DNA) prior to initiating alemtuzumab given the lack of its reversibility once administered[243].

The issues surrounding switching patients with RRMS, particularly if JCV-positive, have been considered by Giovannoni et al. along with practical suggestions for approaches to this strategy[248]. This group advocates a washout period of less than 4 weeks between stopping natalizumab and starting a new DMT in general but has a different strategy for alemtuzumab given the risk of subclinical PML. Using a 'bridging' agent is suggested in order to allow the immune system to reconstitute after natalizumab discontinuation, and hence detect and manage subclinical PML, whilst still providing treatment for the patient's often active MS. Fingolimod is suggested as the bridging agent of choice but MRI and CSF for JCV DNA are routinely undertaken in their centre before starting this. If these investigations do not suggest PML, fingolimod is commenced 2-4 weeks after the last natalizumab infusion and alemtuzumab 6-12 months thereafter. Our data support a switch from natalizumab to alemtuzumab in less than 4 months, without the use of a bridging DMT, to maximise disease-free survival and limit breakthrough disease activity during the switch.

Overall, there is a lack of evidence on the best approach to switching from natalizumab to alemtuzumab and the literature available is based on small cohorts, short follow-up or both. ANSWERS MS provides much needed evidence of the safety and efficacy of switching natalizumab to alemtuzumab using real-world data with long-term follow-up unique to our cohort.

Conclusions

This real-world multicentre retrospective observational study suggests that alemtuzumab is effective when natalizumab fails in patients with active RRMS, and that sequencing these treatments does not increase risk beyond what is expected with alemtuzumab use first-line. Follow-up of almost 3 years for safety outcomes in 79 patients and 4.8 years for efficacy outcomes in 51 patients provides the largest dataset with longitudinal outcomes available worldwide for this increasingly used treatment strategy. Despite sequential high-efficacy DMTs and robust suppression of inflammatory disease activity particularly during the first 2 years of alemtuzumab treatment, our cohort had gradually worsening disability with time, although this plateaued after alemtuzumab treatment. Our data support switching from natalizumab to alemtuzumab directly within at most

4 months to maximise disease-free survival, but would caution to balance this with safety by excluding (subclinical) PML as far as possible before making the switch to an irreversible immunosuppressant. There remains ongoing worldwide interest in this approach to treatment with a prospective study now underway in the United States which may overcome some of the limitations of our study. Until this is reported, however, our data can help guide clinicians to counsel patients on the safety and efficacy of switching from natalizumab to alemtuzumab in active RRMS.

Final conclusions and recommendations

Based on a sound understanding of the natural history of MS and a detailed review of the literature pertaining to DMTs, this thesis has explored current treatment strategies at local, national and international levels and their potential impact on patients now and in the future. Cohort studies detailing the natural history of MS in the pre-treatment era will not be replicated, but provide the foundation of understanding against which the utility of treatments is measured. However, changes in society, healthcare and MS diagnostic definitions, amongst other factors, mean that true comparisons of treatment strategies are possible only between contemporary cohorts.

The lifelong nature of MS and the practicalities of longitudinal interventions and data collection limit the use of gold-standard randomised controlled trials in establishing the utility of treatments in the long-term, against either placebo or each other. The clear short- to medium-term benefits of DMTs in RRMS largely preclude further placebo-controlled trials on both ethical and methodological grounds, with both clinicians and patients opting for active treatment with increasing disease activity. This thesis is therefore part of the growing body of work recognising the necessity and utility of (particularly long-term) real-world observational studies in determining the true value of MS therapeutics. We have provided safety and efficacy data on commonly used DMTs in our centre as well as less common sequencing strategies with highly efficacious treatments using multicentre data. Additionally, we have demonstrated the variability of DMT use within Scotland and the potential impact this can have even within a relatively small cohort.

The pivotal randomised trials of current and emerging therapies in MS have been appraised here and the need for ongoing post-licensing evaluation recognised. The use of DMTs early in the disease course of MS is supported by current literature and our local data, but important safety issues may only emerge once treatments are used on a wider scale outwith research trials. Evaluating the relatively recent introduction of oral therapies for RRMS in our centre, we have found that efficacy outcomes are comparable to that expected from randomised trials, but rates of both physical and biochemical side effects were often higher than expected and should inform decision-making regarding their use.

Using Scottish real-world data, we have established the variability in use of DMTs in RRMS, both in terms of their initiation and escalation. We identified variation in the use of DMTs in Scotland not accounted for by disease characteristics alone, suggesting similar patients are being treated differently for other reasons – this is the first time that this variation has been confirmed scientifically, though it has long been suspected. As outlined, there is clear support in the literature for early DMT initiation in RRMS, raising questions as to why such variation exists in Scotland. We also identified evidence of benefit on relapse activity in early treated patients, in keeping with the expectation from published literature.

Other decisions in MS therapeutics have much less of an evidence-base, specifically sequencing and switching DMTs. With this in mind, we evaluated all patients we could identify in the UK & Ireland with longitudinal follow-up after switching between two highly effective DMTs, natalizumab and alemtuzumab, to clarify the safety and efficacy of this increasingly used approach. Reassuringly, we found no new concerning safety signals, and there were clear benefits in using alemtuzumab where natalizumab failed in terms of inflammatory disease activity. Additionally, we have been able to provide guidance on the switching process, suggesting a shorter switch time has efficacy benefits if balanced against potential risks.

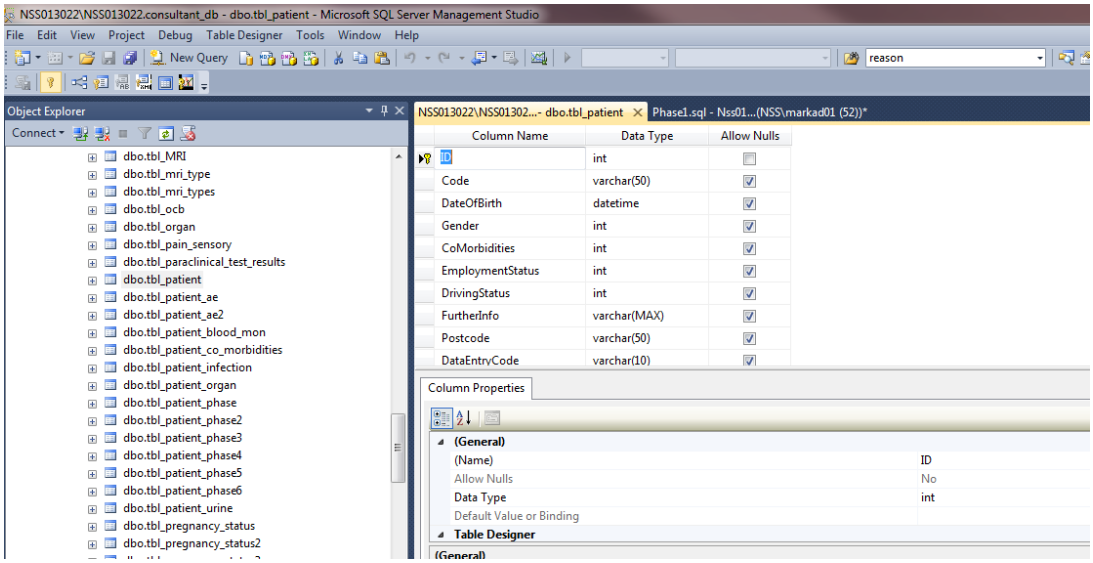
Both ANSWERS and MODERATE are examples of the power of collaboration between centres, in order to pool data and provide more robust generally applicable results than a single centre alone. The limitations common to observational studies were not averted entirely in our studies but we have taken steps to minimise at least some of these where possible, such as statistical matching, to reduce bias. Inherent limitations, particularly unequal follow-up in MODERATE, limited conclusions on the impact of treatment variability. Meaningful long-term outcomes are what really matters in MS and the prospect of future study of both the MODERATE and ANSWERS MS cohorts may provide further data to achieve this goal more fully.

The challenge going forward is to expand the scope and quality of real-world observational data so that robust conclusions can be drawn. International, multicentre, collaborative studies have already provided the beginnings of this in harnessing large datasets to report on safety and efficacy of treatments in a way

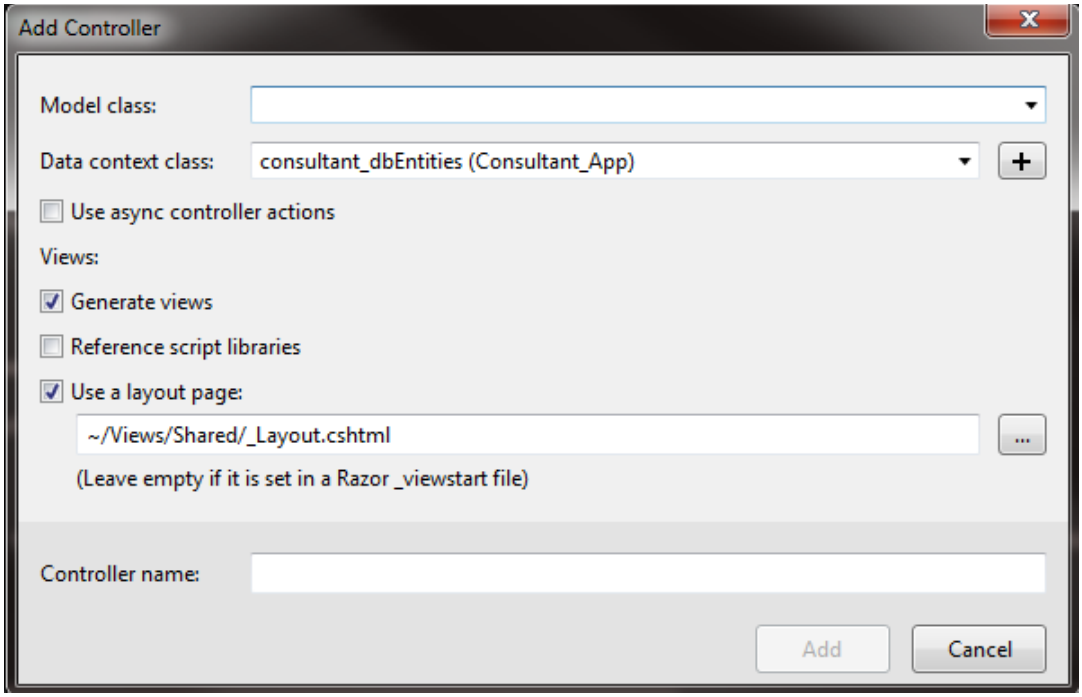
that short-term randomised trials simply cannot. We hope to use the methods and results presented here as a springboard for further pragmatic study of DMTs in MS, to enhance our understanding of the role of current and emerging therapies in improving the lives of people with MS.

Appendices

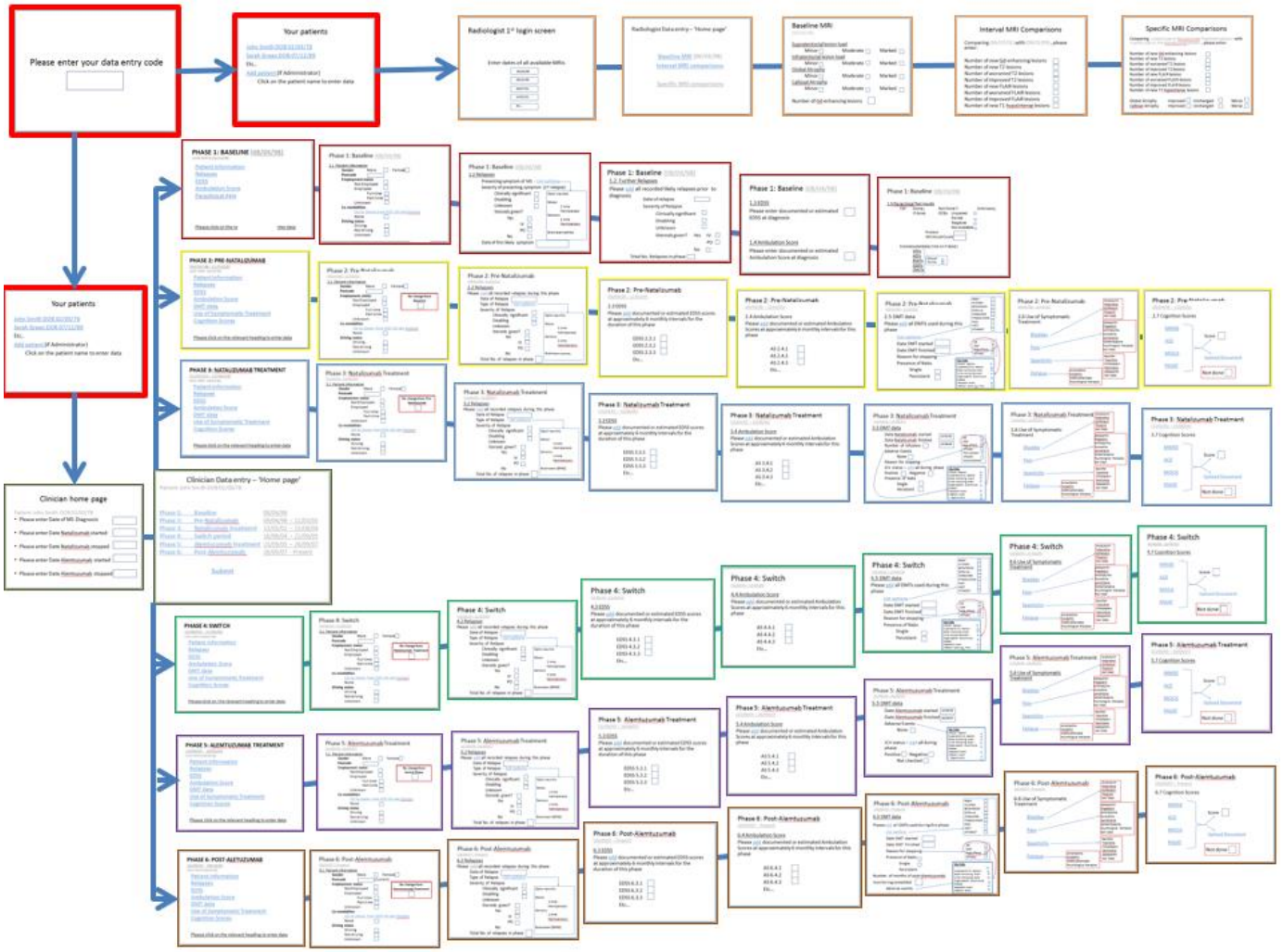
Appendix 1: ANSWERS MS database software - Patient table



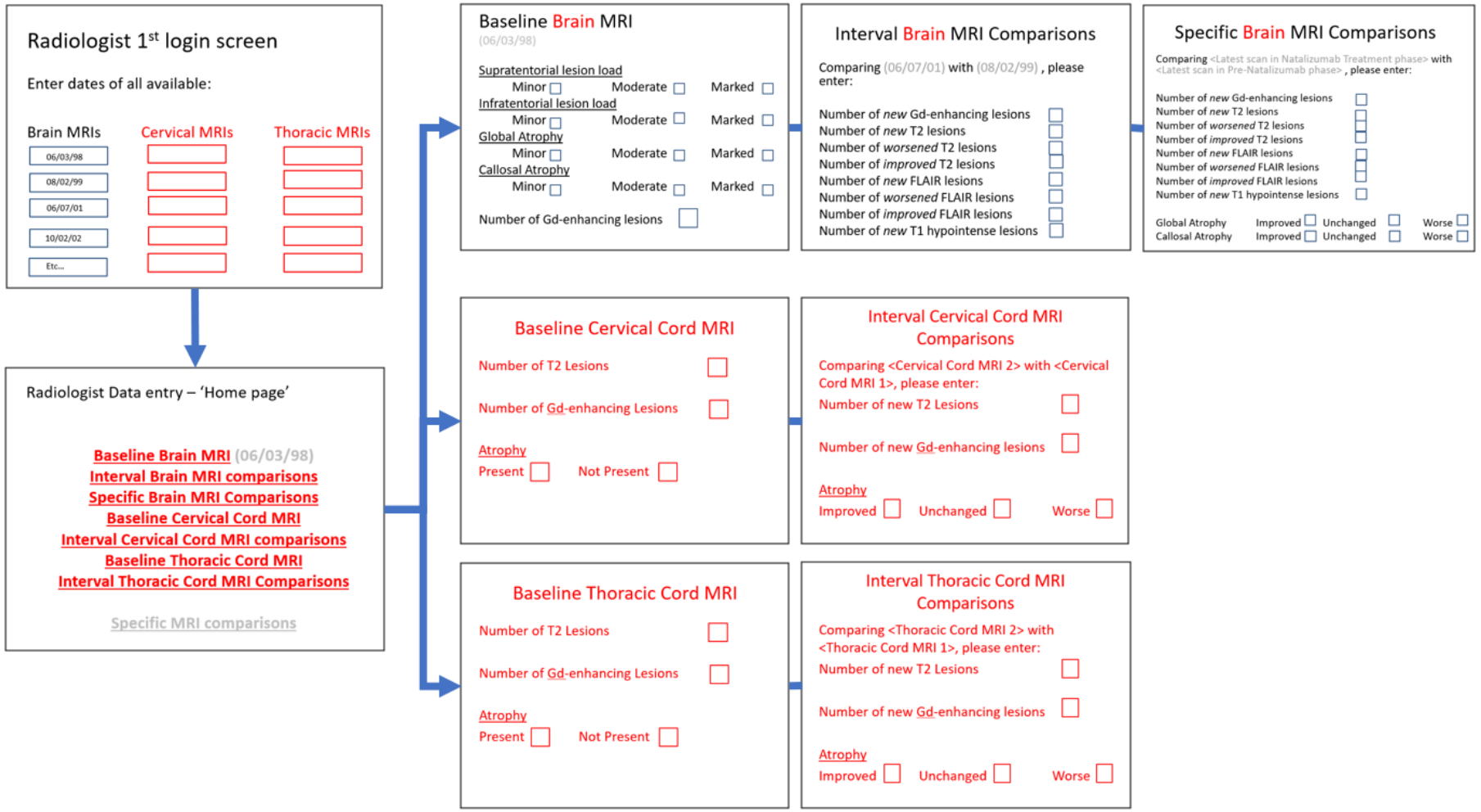
Appendix 2: ANSWERS MS database software - Adding a controller



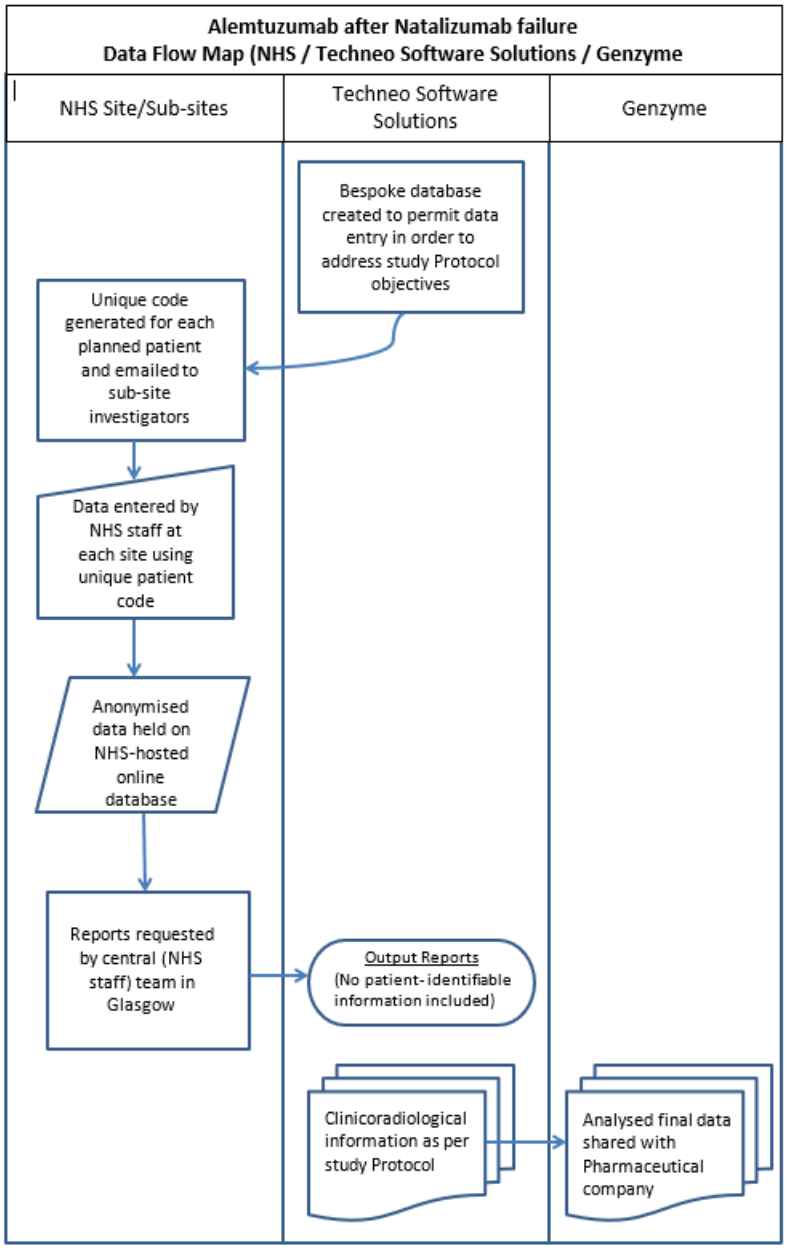
Appendix 3: ANSWERS MS Clinical database outline


Please enter your data entry code
Your patients
Radiologist 1st login screen
Radiologist Data entry – 'Home page'
Baseline MRI
Interval MRI Comparisons
Specific MRI Comparisons
Your patients
Clinician home page
Clinician Data entry – 'Home page'
PHASE 1: BASELINE
Phase 1: Baseline
PHASE 2: PRE-NATALIZUMAB
Phase 2: Pre-Natalizumab
PHASE 3: NATALIZUMAB TREATMENT
Phase 3: Natalizumab Treatment
PHASE 4: SWITCH
Phase 4: Switch
PHASE 5: ALEMTUZUMAB TREATMENT
Phase 5: Alemtuzumab Treatment
PHASE 6: POST-ALETUZUMAB
Phase 6: Post-Alemtuzumab

Appendix 4: ANSWERS MS MRI database outline



Appendix 5: ANSWERS MS Data Flow map



References

1. World Health Organisation. Atlas: Multiple Sclerosis Resources in the World. *Vasa* (2008). doi:ISBN 978 92 4 156375 8

2. Trisolini M, Honeycutt A, Wiener J, L. S. Global Economic Impact of Multiple Sclerosis. *Mult. Scler. Int. Fed.* (2010).

3. Koch-Henriksen, N. & Sørensen, P. S. The changing demographic pattern of multiple sclerosis epidemiology. *Lancet Neurol.* **9**, 520–532 (2010).

4. Pugliatti, M. *et al.* The epidemiology of multiple sclerosis in Europe. *Eur. J. Neurol.* **13**, 700–722 (2006).

5. Lucchinetti, C. *et al.* Heterogeneity of multiple sclerosis lesions: Implications for the pathogenesis of demyelination. *Ann. Neurol.* **47**, 707–717 (2000).

6. König, F. B. *et al.* Persistence of immunopathological and radiological traits in multiple sclerosis. *Arch. Neurol.* **65**, 1527–32 (2008).

7. Metz, I. *et al.* Pathologic heterogeneity persists in early active multiple sclerosis lesions. *Ann. Neurol.* **75**, 728–738 (2014).

8. Keegan, M. *et al.* Relation between humoral pathological changes in multiple sclerosis and response to therapeutic plasma exchange. *Lancet (London, England)* **366**, 579–82 (2005).

9. Pittock, S. J. *et al.* Clinical course, pathological correlations, and outcome of biopsy proved inflammatory demyelinating disease. *J. Neurol. Neurosurg. Psychiatry* **76**, 1693–7 (2005).

10. Goodin, D. S. Magnetic resonance imaging as a surrogate outcome measure of disability in multiple sclerosis: Have we been overly harsh in our assessment? *Ann. Neurol.* **59**, 597–605 (2006).

11. Palace, J. *et al.* 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**, 497–505 (2015).

12. Weinshenker, B. G. *et al*. The natural history of multiple sclerosis: a geographically based study. 2. Predictive value of the early clinical course. *Brain* **112 (Pt 6**, 1419–1428 (1989).

13. Confavreux, C., Vukusic, S. & Adeleine, P. Early clinical predictors and progression of irreversible disability in multiple sclerosis: An amnesic process. *Brain* **126**, 770–782 (2003).

14. Pittock, S. J. *et al*. Disability profile of MS did not change over 10 years in a population-based prevalence cohort. *Neurology* **62**, 601–606 (2004).

15. Tremlett, H., Paty, D. & Devonshire, V. Disability progression in multiple sclerosis is slower than previously reported. *Neurology* **66**, 172–177 (2006).

16. Brown, M. G. *et al*. Estimating typical multiple sclerosis disability progression speed from clinical observations. *PLoS One* **9**, (2014).

17. Leray, E. *et al*. Evidence for a two-stage disability progression in multiple sclerosis. *Brain* **133**, 1900–1913 (2010).

18. Debouverie, M., Pittion-Vouyovitch, S., Louis, S. & Guillemin, F. Natural history of multiple sclerosis in a population-based cohort. *Eur. J. Neurol.* **15**, 916–921 (2008).

19. Debouverie, M., Louis, S., Pittion-Vouyovitch, S., Roederer, T. & Vespignani, H. Multiple sclerosis with a progressive course from onset in Lorraine-Eastern France. *J. Neurol.* **254**, 1370–1375 (2007).

20. Tremlett, H., Zhao, Y., Rieckmann, P. & Hutchinson, M. New perspectives in the natural history of multiple sclerosis. *Neurology* **74**, 2004–2015 (2010).

21. Weinshenker, B. G. *et al*. The natural history of multiple sclerosis: a geographically based study. I. Clinical course and disability. *Brain* **112 (Pt 1**, 133–46 (1989).

22. Poser, C. M. *et al*. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann. Neurol.* **13**, 227–231 (1983).

23. McDonald, W. I. *et al*. Recommended diagnostic criteria for multiple

sclerosis: Guidelines from the International Panel on the Diagnosis of Multiple Sclerosis. *Ann. Neurol.* **50**, 121–127 (2001).

24. Weinshenker, B. G. *et al.* The natural history of multiple sclerosis: a geographically based study. 3. Multivariate analysis of predictive factors and models of outcome. *Brain* **114 (Pt 2**, 1045–1056 (1991).

25. Kurtzke, J. F. A new scale for evaluating disability in multiple sclerosis. *Neurology* **5**, 580–583 (1955).

26. Kurtzke, J. F. Rating neurologic impairment in multiple sclerosis: An expnaded disability status scale (EDSS). *Neurology* **15**, 654–661 (1983).

27. Bonduellle, M. & Albaranes, R. Etude statistique de 145 cas de la sclerose en plaque. *Sem Hop Paris* **38**, 3762–3773 (1962).

28. Minderhoud, J. M., van der Hoeven, J. H. & Prange, a J. Course and prognosis of chronic progressive multiple sclerosis. Results of an epidemiological study. *Acta Neurol. Scand.* **78**, 10–15 (1988).

29. Cottrell, D. A. *et al.* The natural history of multiple sclerosis: a geographically based study. 5. The clinical features and natural history of primary progressive multiple sclerosis. *Brain* **122 (Pt 4**, 625–639 (1999).

30. Kremenchutzky, M. *et al.* The natural history of multiple sclerosis: a geographically based study. 7. Progressive-relapsing and relapsing-progressive multiple sclerosis: a re-evaluation. *Brain* **122 (Pt 1**, 1941–1950 (1999).

31. Kremenchutzky, M., Rice, G. P. A., Baskerville, J., Wingerchuk, D. M. & Ebers, G. C. The natural history of multiple sclerosis: A geographically based study 9: Observations on the progressive phase of the disease. *Brain* **129**, 584–594 (2006).

32. Scalfari, A. *et al.* The natural history of multiple sclerosis, a geographically based study 10: Relapses and long-term disability. *Brain* **133**, 1914–1929 (2010).

33. Tremlett, H., Yousefi, M., Devonshire, V., Rieckmann, P. & Zhao, Y. Impact of multiple sclerosis relapses on progression diminishes with time.

Neurology **73**, 1616–1623 (2009).

34. Lublin, F. D., Baier, M. & Cutter, G. Effect of relapses on development of residual deficit in multiple sclerosis. *Neurology* **61**, 1528–1532 (2003).

35. Kalincik, T. *et al*. Defining reliable disability outcomes in multiple sclerosis. *Brain* **138**, 3287–3298 (2015).

36. Liu, C. & Blumhardt, L. D. Disability outcome measures in therapeutic trials of relapsing-remitting multiple sclerosis: effects of heterogeneity of disease course in placebo cohorts. *J. Neurol. Neurosurg. Psychiatry* **68**, 450–457 (2000).

37. Tremlett, H., Zhao, Y., Joseph, J. & Devonshire, V. Relapses in multiple sclerosis are age- and time-dependent. *J. Neurol. Neurosurg. Psychiatry* **79**, 1368–1374 (2008).

38. Murray, T. J. The history of multiple sclerosis: the changing frame of the disease over the centuries. *J. Neurol. Sci.* **277**, (2009).

39. Brain, R. Critical review: disseminating sclerosis. *Q. J. Med.* **23**, 343 (1930).

40. Miller, H., Newewll, D. & Ridley, A. Multiple Sclerosis: Trials of maintenance treatment with prednisolone and soluble aspirin. *Lancet* **1**, 127–9 (1961).

41. Brown, J. R. *et al*. The design of clinical studies to assess therapeutic efficacy in multiple sclerosis. *Neurology* **29**, 3–23 (1979).

42. Schumacher, G., Beebe, G., Kibler, R. & Kurland, L. Problems of experimental trials of therapy in multiple sclerosis. *Ann. N. Y. Acad. Sci.* **122**, 552–68 (1965).

43. Goodin, D. S. Disease-modifying therapy in MS: a critical review of the literature. Part I: Analysis of clinical trial errors. *J. Neurol.* **251 Suppl**, v3–v11 (2004).

44. Montalban, X. Review of methodological issues of clinical trials in multiple sclerosis. *J. Neurol. Sci.* **311**, (2011).

45. Goodin, D. S. Disease-modifying therapy in MS: A critical review of the literature. Part II: Assessing efficacy and dose-response. in *Journal of Neurology, Supplement* **251**, (2004).

46. Goodman, S. N. p Values, Hypothesis Tests, and Likelihood: Implications for Epidemiology of a Neglected Historical Debate. *Am. J. Epidemiol.* **137**, 485–501 (1993).

47. Panitch, H. *et al.* Randomized, comparative study of interferon beta-1a treatment regimens in MS: The EVIDENCE Trial. *Neurology* **59**, 1496–1506 (2002).

48. Van Walraven, C., Davis, D., Forster, A. J. & Wells, G. A. Time-dependent bias was common in survival analyses published in leading clinical journals. *J. Clin. Epidemiol.* **57**, 672–682 (2004).

49. Cohen, J. a *et al.* Alemtuzumab versus interferon beta 1a as first-line treatment for patients with relapsing-remitting multiple sclerosis: a randomised controlled phase 3 trial. *Lancet* **380**, 1819–28 (2012).

50. Dahdaleh, D. & Sharrack, B. We can compare the relative efficacy of multiple sclerosis medications by examining the results of indpendent clinical trials: Yes. *Mult. Scler. J.* **21**, 35–36 (2015).

51. Ciccarelli, O. & Chataway, J. We can compare the relative efficacy of multiple sclerosis medications by examining the results of indpendent clinical trials: No. *Mult. Scler. J.* **21**, 37–38 (2015).

52. Tramacere, I., Del Giovane, C., Salanti, G., D'Amico, R. & Filippini, G. Immunomodulators and immunosuppressants for multiple sclerosis: a network meta-analysis. *Cochrane database Syst. Rev.* CD011381 (2015). doi:10.1002/14651858.CD008933.pub2

53. Del Santo, F., Maratea, D., Fadda, V., Trippoli, S. & Messori, A. Treatments for relapsing-remitting multiple sclerosis: Summarising current information by network meta-analysis. *Eur. J. Clin. Pharmacol.* **68**, 441–448 (2012).

54. Benson, K. & Hartz, A. J. A comparison of observational studies and

randomized, controlled trials. *N. Engl. J. Med.* **342**, 1878–86 (2000).

55. Montalban, X. Review of methodological issues of clinical trials in multiple sclerosis. *J. Neurol. Sci.* **311**, S35–S42 (2011).

56. Fogarty, E., Schmitz, S., Tubridy, N., Walsh, C. & Barry, M. Comparaitve efficacy of disease-modifying therapies for patients with relapsing remitting multiple sclerosis: systematic review and network meta-analysis. *Mult. Scler. Relat. Disord.* **9**, 23–30 (2016).

57. Jacobs, L. D. *et al.* Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group (MSCRG). *Ann. Neurol.* **39**, 285–294 (1996).

58. The IFNB Multiple Sclerosis Study Group. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. I. Clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. *Neurology* **43**, 655–661 (1993).

59. Ebers, G. C. Randomised double-blind placebo-controlled study of interferon beta-1a in relapsing/remitting multiple sclerosis. PRISMS (Prevention of Relapses and Disability by Interferon beta-1a Subcutaneously in Multiple Sclerosis) Study Group. *Lancet* **352**, 1498–504 (1998).

60. Calabresi, P. A. *et al.* Pegylated interferon beta-1a for relapsing-remitting multiple sclerosis (ADVANCE): A randomised, phase 3, double-blind study. *Lancet Neurol.* **13**, 657–665 (2014).

61. Johnson, K. P. *et al.* Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multicenter, double-blind placebo-controlled trial. The Copolymer 1 Multiple Sclerosis Study Group. *Neurology* **45**, 1268–1276 (1995).

62. O'Connor, P. *et al.* Randomized trial of oral teriflunomide for relapsing multiple sclerosis. *N. Engl. J. Med.* **365**, 1293–303 (2011).

63. Confavreux, C. *et al.* Oral teriflunomide for patients with relapsing multiple sclerosis (TOWER): A randomised, double-blind, placebo-

controlled, phase 3 trial. *Lancet Neurol.* **13**, 247–256 (2014).

64. Gold, R. *et al.* Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis. *N. Engl. J. Med.* **367**, 1098–107 (2012).

65. Fox, R. J. *et al.* Placebo-controlled phase 3 study of oral BG-12 or glatiramer in multiple sclerosis. *N. Engl. J. Med.* **367**, 1087–97 (2012).

66. Kappos, L. *et al.* A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *N. Engl. J. Med.* **362**, 387–401 (2010).

67. Calabresi, P. a *et al.* Safety and efficacy of fingolimod in patients with relapsing-remitting multiple sclerosis (FREEDOMS II): a double-blind, randomised, placebo-controlled, phase 3 trial. *Lancet Neurol.* **13**, 545–56 (2014).

68. Polman, C. H. *et al.* A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *N. Engl. J. Med.* **354**, 899–910 (2006).

69. Cohen, J. A. *et al.* Alemtuzumab versus interferon beta 1a as first-line treatment for patients with relapsing-remitting multiple sclerosis: a randomised controlled phase 3 trial. *Lancet* **380**, 1819–1828 (2012).

70. Coles, A. J. *et al.* Alemtuzumab for patients with relapsing multiple sclerosis after disease-modifying therapy: A randomised controlled phase 3 trial. *Lancet* **380**, 1829–1839 (2012).

71. Hauser, S. L., Bar-Or, A., Comi, G. & Giovannoni, G. Ocrelizumab versus Interferon Beta-1a in Relapsing Multiple Sclerosis. *N. Engl. J. Med.* **376**, 221–34 (2017).

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

73. Comi, G. *et al.* Placebo-controlled trial of oral laquinimod for multiple sclerosis. Supplementary Appendix. *N. Engl. J. Med.* **366**, 1000–9 (2012).

74. Gold, R. *et al.* Daclizumab high-yield process in relapsing-remitting multiple sclerosis (SELECT): A randomised, double-blind, placebo-controlled trial. *Lancet* **381**, 2167–75 (2013).

75. European Medicines Agency. *EMA review of Zinbryta confirms medicine's risks outweigh its benefits*. (2018).

76. Meuth, S. *et al*. Therapeutic approaches to multiple sclerosis: an update on failed, interrupted or inconclusive trials of neuroprotective and alternative treatment strategies. *Biodrugs* **24**, 317–330 (2010).

77. Cadavid, D. *et al*. Efficacy and analysis of opicinumab in relapsing remitting multiple sclerosis: The phase 2b SYNERGY trial. *Mult. Scler.* **22**, 66 (2016).

78. Raftopoulos, R. *et al*. Phenytoin for neuroprotection in patients with acute optic neuritis: a randomised, placebo-controlled, phase 2 trial. *Lancet Neurol.* **15**, 259–69 (2016).

79. Metz, L. M. *et al*. Trial of minocycline in a clinically isolated syndrome of multiple sclerosis. *N. Engl. J. Med.* **376**, 2122–33 (2017).

80. Green, A. J. *et al*. Clemastine fumarate as a remyelinating therapy for multiple sclerosis (ReBUILD): a randomised, controlled, double-blind, crossover trial. *Lancet* **390**, 2481–89 (2017).

81. Montalban, X., Hauser, S. L., Kappos, L., Arnold, D. L. & Bar-Or, A. Ocrelizumab versus placebo in primary progressive multiple sclerosis. *N. Engl. J. Med.* **376**, 209–20 (2017).

82. Kappos, L. *et al*. Siponimod versus placebo in secondary progressive multiple sclerosis (EXPAND): a double-blind, randomised, phase 3 study. *Lancet* **391**, 1263–73 (2018).

83. Tourbah, A. *et al*. MD1003 (high-dose biotin) for the treatment of progressive multiple sclerosis: A randomised, double-blind, placebo-controlled study. *Mult. Scler.* **22**, 1719–1731 (2016).

84. Giovannoni, G. *et al*. Is multiple sclerosis a length-dependent central axonopathy? The case for therapeutic lag and the asynchronous progressive MS hypotheses. *Mult. Scler. Relat. Disord.* **12**, 70–78 (2017).

85. Marrie, R. A., Cohen, J. & Stuve, O. A systematic review of the incidence and prevalance of comorbidity in multiple sclerosis. *Mult. Scler.* 1–19

(2015). doi:10.1177/ 1352458514564491

86. Thompson, P. A., Banwell, M. P. B., Barkhof, P. F. & Et.al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* **17**, 162–173 (2018).

87. Okuda, D. T. *et al.* Incidental MRI anomalies suggestive of multiple sclerosis: The radiologically isolated syndrome. *Neurology* **72**, 800–805 (2009).

88. Granberg, T., Martola, J., Kristoffersen-Wiberg, M., Aspelin, P. & Fredrikson, S. Radiologically isolated syndrome - incidental magnetic resonance imaging findings suggestive of multiple sclerosis, a systematic review. *Mult. Scler. J.* **19**, 271–80 (2013).

89. Sorensen, P. S. Early-stage multiple sclerosis: What are the treatment options? *Drugs* **64**, 2021–2029 (2004).

90. Tuohy, V. K. *et al.* The epitope spreading cascade during progression of experimental autoimmune encephalomyelitis and multiple sclerosis. *Immunol. Rev.* **164**, 93–100 (1998).

91. Trapp, B. D. *et al.* Axonal transection in the lesions of multiple sclerosis. *N. Engl. J. Med.* **338**, 278–285 (1998).

92. Kuhlmann, T., Lingfeld, G., Bitsch, A., Schuchardt, J. & Brück, W. Acute axonal damage in multiple sclerosis is most extensive in early disease stages and decreases over time. *Brain* **125**, 2202–2212 (2002).

93. Brex, P. A. *et al.* Detection of ventricular enlargement in patients at the earliest clinical stage of MS. *Neurology* **54**, 1689–1691 (2000).

94. Audoin, B. *et al.* Onset and underpinnings of white matter atrophy at the very early stage of multiple sclerosis - a two-year longitudinal MRI/MRSI study of corpus callosum. *Mult. Scler.* **13**, 41–51 (2007).

95. Liu, C., Edwards, S., Gong, Q., Roberts, N. & Blumhardt, L. D. Three dimensional MRI estimates of brain and spinal cord atrophy in multiple sclerosis. *J Neurol Neurosurg Psychiatry* **66**, 323–330 (1999).

96. Schlaeger, R. *et al*. Spinal cord gray matter atrophy correlates with multiple sclerosis disability. *Annals of Neurology* (2014). doi:10.1002/ana.24241

97. Fisher, E. *et al*. Eight-year follow-up study of brain atrophy in patients with MS. *Neurology* **59**, 1412–1420 (2002).

98. Sormani, M. P., Arnold, D. L. & De Stefano, N. Treatment effect on brain atrophy correlates with treatment effect on disability in multiple sclerosis. *Ann. Neurol.* **75**, 43–49 (2014).

99. Coles, A. J. *et al*. The window of therapeutic opportunity in multiple sclerosis: Evidence from monoclonal antibody therapy. *J. Neurol.* **253**, 98–108 (2006).

100. Lorscheider, J. *et al*. Defining secondary progressive multiple sclerosis. *Brain* **139**, 2395–2405 (2016).

101. Lublin, F. D. *et al*. Defining the clinical course of multiple sclerosis: The 2013 revisions. *Neurology* **83**, 278–286 (2014).

102. Nicoletti, F. *et al*. Elevated serum levels of interleukin-12 in chronic progressive multiple sclerosis. *J. Neuroimmunol.* **70**, 87–90 (1996).

103. Jacobs, L. D. *et al*. Intramuscular interferon beta-1a therapy initiated during a first demyelinating event in multiple sclerosis. CHAMPS Study Group. *N. Engl. J. Med.* **343**, 898–904 (2000).

104. Comi, G. *et al*. Effect of early interferon treatment on conversion to definite multiple sclerosis: A randomised study. *Lancet* **357**, 1576–1582 (2001).

105. Kappos, L. *et al*. Treatment with interferon beta-1b delays conversion to clinically definite and McDonald MS in patients with clinically isolated syndromes. *Neurology* **67**, 1242–1249 (2006).

106. Comi, G. *et al*. Effect of glatiramer acetate on conversion to clinically definite multiple sclerosis in patients with clinically isolated syndrome (PreCISe study): a randomised, double-blind, placebo-controlled trial. *Lancet* **374**, 1503–1511 (2009).

107. Kappos, L. *et al*. Long-term subcutaneous interferon beta-1a therapy in patients with relapsing-remitting MS. *Neurology* **67**, 944–953 (2006).

108. Goodin, D. S. *et al*. Relationship between early clinical characteristics and long term disability outcomes: 16 year cohort study (follow-up) of the pivotal interferon ß-1b trial in multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **83**, 282–7 (2012).

109. Goodin, D. S. *et al*. Survival in MS A randomized cohort study 21 years after the start of the pivotal IFN-1b trial. *Neurology* **78**, 1315–1322 (2012).

110. Chalmer, T. A. *et al*. Early versus later treatment start in multiple sclerosis: a register-based cohort study. *Eur. J. Neurol.* **25**, 1262–1270 (2018).

111. Trojano, M. *et al*. New natural history of interferon-beta-treated relapsing multiple sclerosis. *Ann. Neurol.* **61**, 300–306 (2007).

112. Freedman, M. S. Induction vs. escalation of therapy for relapsing Multiple Sclerosis: The evidence. *Neurol. Sci.* **29**, (2008).

113. Rieckmann, P. Concepts of induction and escalation therapy in multiple sclerosis. *J. Neurol. Sci.* **277**, (2009).

114. Coles, A. J. *et al*. Pulsed monoclonal antibody treatment and autoimmune thyroid disease in multiple sclerosis. *Lancet* **354**, 1691–1695 (1999).

115. Jayne, D. Challenges in the management of microscopic polyangiitis: past, present and future. *Curr. Opin. Rheumatol.* **20**, 3–9 (2008).

116. Pui, C.-H. & Evans, W. E. Treatment of acute lymphoblastic leukemia. *N. Engl. J. Med.* **354**, 166–178 (2006).

117. Hauser, S. *et al*. B-Cell Depletion with Rituximab in Relapsing–Remitting Multiple Sclerosis. *N. Engl. J. Med.* **358**, 676–88 (2008).

118. Salzer, J. *et al*. Rituximab in multiple sclerosis. *Neurology* **87**, 2074–2081 (2016).

119. Montalbán, X., Hauser, S. L., Kappos, L. & Arnold, D. L. Ocrelizumab versus Placebo in Primary Progressive Multiple Sclerosis. *N. Engl. J. Med.*

376, 209–20 (2017).

120. Coles, A. J. *et al*. Alemtuzumab for patients with relapsing multiple sclerosis after disease-modifying therapy: a randomised controlled phase 3 trial. *Lancet* **380**, 1829–39 (2012).

121. Rudick, R. A. *et al*. Natalizumab plus Interferon Beta-1a for Relapsing Multiple Sclerosis. *N. Engl. J. Med.* **354**, 911–923 (2006).

122. Goodman, A. D. *et al*. Glance: Results of a phase 2, randomized, double-blind, placebo-controlled study. *Neurology* **72**, 806–812 (2009).

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

124. Vermersch, P. *et al*. Teriflunomide versus subcutaneous interferon beta-1a in patients with relapsing multiple sclerosis: a randomised, controlled phase 3 trial. *Mult. Scler.* **20**, 705–716 (2014).

125. Durelli, L. *et al*. Every-other-day interferon beta-1b versus once-weekly interferon beta-1a for multiple sclerosis: Results of a 2-year prospective randomised multicentre study (INCOMIN). *Lancet* **359**, 1453–1460 (2002).

126. Mikol, D., Barkhof, F., Chang, P. & Coyle, P. Comparison of subcutaneous interferon beta-1a with glatiramer acetate in patients with relapsng multiple sclerosis (the REbif vs Glatiramer Acetate in Relapsing MS Disease [REGARD] study): a multicentre, randomsed, parallel,open-label trial. *Lancet Neurol.* **7**, 903–14 (2008).

127. O'Connor, P. *et al*. 250 Microg or 500 Microg Interferon Beta-1B Versus 20 Mg Glatiramer Acetate in Relapsing-Remitting Multiple Sclerosis: a Prospective, Randomised, Multicentre Study. *Lancet Neurol.* **8**, 889–97 (2009).

128. MHRA. Fingolimod (Gilenya▼): risks of progressive multifocal leukoencephalopathy, basal-cell carcinoma, and opportunistic infections. Available at: https://www.gov.uk/drug-safety-update/fingolimod-gilenya-risks-of-progressive-multifocal-leukoencephalopathy-basal-cell-carcinoma-and-opportunistic-infections.

129. MHRA. Dimethyl fumarate (Tecfidera): updated advice on risk of progressive multifocal leukoencephalopathy. Available at: https://www.gov.uk/drug-safety-update/dimethyl-fumarate-tecfidera-updated-advice-on-risk-of-progressive-multifocal-leukoencephalopathy.

130. Rio, J. Any evident MRI T2 lesion activity should guide change of therapy in multiple sclerosis: No. *Mult. Scler. J.* **21**, 132–133 (2015).

131. Giovannoni, G. Any evident MRI T2 lesion activity should guide change of therapy in multiple sclerosis: Yes. *Mutiple Scler. J.* **21**, 124–126 (2015).

132. Bermel, R. A. *et al.* Predictors of long-term outcome in multiple sclerosis patients treated with interferon beta. *Ann. Neurol.* **73**, 95–103 (2013).

133. Sormani, M. P. & De Stefano, N. Defining and scoring response to IFN-β in multiple sclerosis. *Nat. Rev. Neurol.* **9**, 504–12 (2013).

134. Horakova, D. *et al.* Early predictors of non-response to interferon in multiple sclerosis. *Acta Neurol. Scand.* **126**, 390–397 (2012).

135. Prosperini, L. *et al.* Interferon beta failure predicted by EMA criteria or isolated MRI activity in multiple sclerosis. *Mult. Scler.* **20**, 566–76 (2014).

136. Sormani, M. P. *et al.* Assessing response to interferon-Beta in a multicenter dataset of patients with MS. *Neurology* **87**, 134–140 (2016).

137. Freedman, M. S. *et al.* Treatment optimization in MS: Canadian MS Working Group updated recommendations. *Can. J. Neurol. Sci.* **40**, 307–23 (2013).

138. Sormani, M. P. *et al.* Scoring treatment response in patients with relapsing multiple sclerosis. *Mult. Scler.* **19**, 605–12 (2013).

139. Ebers, G. C. Randomised double-blind placebo-controlled study of interferon β-1a in relapsing/remitting multiple sclerosis. *The Lancet* **352**, 1498–1504 (1998).

140. Jacobs, L. *et al.* Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. *Ann. Neurol.* **39**, 285–294 (1996).

141. Kalincik, T. *et al.* Interferon, azathioprine and corticosteroids in multiple

sclerosis: 6-year follow-up of the ASA cohort. *Clin. Neurol. Neurosurg.* **114**, 940–946 (2012).

142. Davis, M., Ashtamker, N., Steinerman, J. & Knappertz, V. Time course of glatiramer acetate efficacy in patients with RRMS in the GALA study. *Neurol. Neuroimmunol. Neuroinflammation* **4**, e327 (2017).

143. Calabresi, P. a *et al.* The incidence and significance of anti-natalizumab antibodies. *Neurology* **69**, 1391–1403 (2007).

144. Prosperini, L. *et al.* Escalation to natalizumab or switching among immunomodulators in relapsing multiple sclerosis. *Mult. Scler. J.* **18**, 64–71 (2012).

145. Spelman, T. *et al.* Comparative efficacy of switching to natalizumab in active multiple sclerosis. *Ann. Clin. Transl. Neurol.* **2**, 373–387 (2015).

146. D'Amico, E., Leone, C., Zanghi, A., Fermo, S. Lo & Patti, F. Lateral and escalation therapy in relapsing-remitting multiple sclerosis: a comparative study. *J. Neurol.* **263**, 1802–1809 (2016).

147. Spelman, T. *et al.* Risk of early relapse following the switch from injectables to oral agents for multiple sclerosis. *Eur. J. Neurol.* **23**, 729–736 (2016).

148. Trojano, M., Tintoré, M., Montalbán, X. & Hillert, J. Treatment decisions in multipe sclerosis- insights from real-world observational studies. *Nat. Rev. Neurol.* **13**, 105–118 (2017).

149. Ingwersen, J. *et al.* Fingolimod in multiple sclerosis: Mechanisms of action and clinical efficacy. *Clinical Immunology* **142**, 15–24 (2012).

150. Subei, A. M. & Cohen, J. A. Sphingosine 1-Phosphate Receptor Modulators in Multiple Sclerosis. *CNS Drugs* **29**, 565–575 (2015).

151. Zarbin, M. A. *et al.* Ophthalmic evaluations in clinical studies of fingolimod (FTY720) in multiple sclerosis. *Ophthalmology* **120**, 1432–1439 (2013).

152. Novartis Pharmaceuticals UK Ltd. Gilenya 0.5mg hard capsules - Summary of Product Charcateristics. Available at:

https://www.medicines.org.uk/EMC/medicine/24443/SPC/Gilenya+0.5mg+hard+capsules.

153. Scottish Medicines Consortium. Fingolimod (as hydrochloride), 0.5mg hard capsules (Gilenya). (2012). Available at: https://www.scottishmedicines.org.uk/files/advice/fingolimod_Gilenya_RESUBMISSION_FINAL_August_2012_amended_30.08.12.pdf.

154. Rosenkranz, T., Novas, M. & Terborg, C. PML in a Patient with Lymphocytopenia Treated with Dimethyl Fumarate. *N. Engl. J. Med.* **372**, 1476–1478 (2015).

155. Gyang, T. V., Hamel, J., Goodman, A. D., Gross, R. A. & Samkoff, L. Fingolimod-associated PML in a patient with prior immunosuppression. *Neurology* **86**, 1843–45 (2016).

156. Fox, R. J. *et al.* Characterizing absolute lymphocyte count profiles in dimethyl fumarate'treated patients with MS Patient management considerations. *Neurol. Clin. Pract.* **6**, 220–229 (2016).

157. Hunt, D., Kavanagh, D., Drummond, I. & Weller, B. Thrombotic microangiopathy associated with interferon beta. *N. Engl. J. Med.* **370**, 1270 [Letter] (2014).

158. Río, J. *et al.* Measures in the first year of therapy predict the response to interferon ß in MS. *Mult. Scler.* **15**, 848–853 (2009).

159. Spelman, T. *et al.* Comparative efficacy of first-line natalizumab vs IFN-ß or glatiramer acetate in relapsing MS. *Neurol. Clin. Pract.* **6**, 102–115 (2016).

160. Richards, R. Interferon beta in multiple sclerosis. *BMJ* **313**, 1159 (1996).

161. Taylor, D. Funding medicines for people with multiple sclerosis. *BMJ* **323**, 1379–80 (2001).

162. Richards, R. The multiple sclerosis risk sharing scheme - Plus ca change, plus ca meme chose. *BMJ* **340**, 2882 (2010).

163. Rafferty, J. Multiple sclerosis risk sharing scheme: a costly failure. *BMJ*

340, 1672 (2010).

164. Confavreux, C. & Vukusic, S. Natural history of multiple sclerosis: a unifying concept. *Brain* **129**, 606–616 (2006).

165. Casserly, C. & Ebers, G. Relapses do not matter in relation to long-term disability: Yes. *Mult. Scler. J.* **17**, 1412–1414 (2011).

166. Ebers, G. C., Heigenhauser, L., Daumer, M., Lederer, C. & Noseworthy, J. H. Disability as an outcome in MS clinical trials. *Neurology* **71**, 624–631 (2008).

167. Shirani, A. *et al.* Association between use of interferon beta and progression of disability in patients with relapsing-remitting multiple sclerosis. *JAMA* **308**, 247–56 (2012).

168. Drulovic, J. *et al.* Interferon-beta and disability progression in relapsing-remitting multiple sclerosis. *Clin. Neurol. Neurosurg.* **115**, (2013).

169. Simon, J. H. *et al.* A longitudinal study of brain atrophy in relapsing multiple sclerosis. *Neurology* **53**, 139–139 (1999).

170. Zipoli, V. *et al. Cognitive impairment predicts conversion to multiple sclerosis in clinically isolated syndromes. Multiple sclerosis (Houndmills, Basingstoke, England)* **16**, (2010).

171. Popescu, V. *et al.* Brain atrophy and lesion load predict long term disability in multiple sclerosis. *J Neurol Neurosurg Psychiatry* **84**, 1082–1091 (2013).

172. Cohen, J. et al. Fingolimod effect on brain atrophy and clinical/MRI correlations in three phase 3 studies - TRANSFORMS, FREEDOMS and FREEDOMS II. in *AAN* (2013).

173. Coles, A. J. Alemtuzumab slows brain volume loss over 4 years despite most relapsing-remitting multiple sclerosis patients not receiving treatment for 3 years. *Neurology* **84**, Supplement P7.263 (2015).

174. Portaccio, E. *et al.* Natalizumab may reduce cognitive changes and brain atrophy rate in relapsing-remitting multiple sclerosis: A prospective, non-

randomized pilot study. *Eur. J. Neurol.* **20**, 986–990 (2013).

175. Rudick, R. A., Fisher, E., Lee, J. C., Simon, J. & Jacobs, L. *Use of the brain parenchymal fraction to measure whole brain atrophy in relapsing-remitting MS. Multiple Sclerosis Collaborative Research Group. Neurology* **53**, (1999).

176. Fisniku, L. K. *et al*. Disability and T2 MRI lesions: A 20-year follow-up of patients with relapse onset of multiple sclerosis. *Brain* **131**, 808–817 (2008).

177. Pittock, S. J. *et al*. Clinical implications of benign multiple sclerosis: A 20-year population-based follow-up study. *Ann. Neurol.* **56**, 303–306 (2004).

178. Amato, M. P. & Portaccio, E. Truly benign multiple sclerosis is rare. *Mult. Scler. J.* **18**, 13–14 (2012).

179. Gajofatto, A., Turatti, M., Bianchi, M. & Forlivesi, S. Benign multiple sclerosis: physical and cognitive impairment follow distinct evoilutions. *Acta Neurol. Scand.* doi:10.1111/ane.1242

180. Degenhardt, A., Ramagopalan, S. V, Scalfari, A. & Ebers, G. C. Clinical prognostic factors in multiple sclerosis: a natural history review. *Nat. Rev. Neurol.* **5**, 672–682 (2009).

181. Langer-Gould, A. *et al*. Clinical and demographic predictors of long-term disability in patients with relapsing-remitting multiple sclerosis: a systematic review. *Arch. Neurol.* **63**, 1686–1691 (2006).

182. Bergamaschi R, Montomoli C, M. G. BREMSO: a simple clinical score for the early prediction of multiple sclerosis long-term evolution. *ECTRIMS* (2013).

183. Polman, C. H. *et al*. Diagnostic criteria for multiple sclerosis: 2010 Revisions to the McDonald criteria. *Ann. Neurol.* **69**, 292–302 (2011).

184. Freedman, M. S. Do not treat from CIS onset: evaluate disease course and prognosis first: No (treat). *Mult. Scler. J.* **18**, 394–395 (2012).

185. Bunyan, R. F. & Pittock, S. J. Do not treat from CIS onset: evaluate disease course and prognosis first: Yes. *Mult. Scler. J.* **18**, 391–393 (2012).

186. Langer-Gould, A., Atlas, S. W., Green, A. J., Bollen, A. W. & Pelletier, D. *Progressive multifocal leukoencephalopathy in a patient treated with natalizumab*. *The New England journal of medicine* **353**, 375–381 (2005).

187. Kleinschmidt-DeMasters, B. K. & Tyler, K. L. *Progressive multifocal leukoencephalopathy complicating treatment with natalizumab and interferon beta-1a for multiple sclerosis*. *The New England journal of medicine* **353**, 369–374 (2005).

188. Bloomgren, G. *et al*. Risk of Natalizumab-Associated Progressive Multifocal Leukoencephalopathy. *New England Journal of Medicine* **366**, 1870–1880 (2012).

189. Kalincik, T. *et al*. Sex as a determinant of relapse incidence and progressive course of multiple sclerosis. *Brain* **136**, 3609–17 (2013).

190. MS Society. *A lottery of treatment and cae - MS services across Scotland and the UK*. (2013).

191. Giovannoni, G. *et al*. Effect of daclizumab high-yield process in patients with highly active relapsing-remitting multiple sclerosis. *J. Neurol.* **261**, 316–323 (2014).

192. Trojano, M. *et al*. Real-life impact of early interferon beta therapy in relapsing multiple sclerosis. *Ann. Neurol.* **66**, 513–520 (2009).

193. Kalincik, T., Horakova, D., Spelman, T. & Jokubaitis, V. Switch to Natalizumab versus Fingolimod in Active Relapsing-Remitting Multiple Sclerosis. *Ann. Neurol.* **77**, 425–435 (2015).

194. Grassi, M., Enea, D., Marchetti, R. & Caricati, A. Combined counseling and bupropion therapy for smoking cessation: identification of outcome predictors. *Drug Dev. Res.* **67**, 208–216 (2006).

195. Connors, A. F. *et al*. The effectiveness of right heart catheterization in the initial care of critically ill patients. SUPPORT Investigators. *JAMA* **276**, 889–897 (1996).

196. MSIF. *Atlas of MS 2013: Mapping Multiple Sclerosis Around the World*. *Multiple Sclerosis International Federation* (2013).

197. Wilsden, T., Barron, A., Mitchell-Heggs, A. & Ginoza, S. *Access to medicines for multiple sclerosis: Challenges and Opportunities*. (2014).

198. Marziniak, M. *et al*. Variations in multiple sclerosis practice within Europe - Is it time for a new treatment guideline? *Mult. Scler. Relat. Disord.* **8**, 35–44 (2016).

199. Swedish National Board of Health and Welfare. *Quality and efficiency in Swedish Health Care - Regional Comparisosn 2012*. (2012).

200. Cameron, E. *et al*. Factors influencing multiple sclerosis disease-modifying treatment prescribing decisions in the United Kingdom: A qualitative interview study. *Mult. Scler. Relat. Disord.* **27**, 378–382 (2019).

201. Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M. & Woloshin, S. Helping doctors and patients make sense of health statistics. *Psychol. Sci. Public Interes. Suppl.* **8**, 53 (2007).

202. Deber, R. B., Kraetschmer, N., Urowitz, S. & Sharpe, N. Do people want to be autonomous patients? Preferred roles in treatment decision-making in several patient populations. *Heal. Expect.* **10**, 248–258 (2007).

203. Heesen, C. *et al*. Benefit-risk perception of natalizumab therapy in neurologists and a large cohort of multiple sclerosis patients. *J. Neurol. Sci.* **376**, 181–190 (2017).

204. Reen, G. K., Silber, E. & Langdon, D. W. Multiple sclerosis patients' understanding and preferences for risks and benefits of disease-modifying drugs: A systematic review. *J. Neurol. Sci.* **375**, 107–122 (2017).

205. Col, N. F. *et al*. Whose Preferences Matter? A Patient-Centered Approach for Eliciting Treatment Goals. *Med. Decis. Mak.* **38**, 44–55 (2017).

206. Kremenchutzky, M. & Walt, L. Perceptions of health status in multiple sclerosis patients and their doctors. *Can. J. Neurol. Sci.* **40**, 210 (2013).

207. Brown, J. W. L., Coles, A. J., Horakova, D., Havrdova, E. & Robertson, N. P. Association of initial disease-modifying therapy with later conversion to secondary progressive multiple sclerosis. *JAMA* **321**, 175–187 (2019).

208. Lunt, M. Selecting an appropriate caliper can be essential for achieving good balance with propensity score matching. *Am. J. Epidemiol.* **179**, 226 (2014).

209. Austin, P. C. & Stuart, E. A. The performance of inverse probability of treatment weighting and full matching on the propensity score in the presence of model misspecification when estimating the effect of treatment on survival outcomes. *Stat. Methods Med. Res.* **26**, 1654 (2017).

210. DeMaria, A. Lies, damned lies, and statistics. *J. Am. Coll. Cardiol.* **52**, 1430 (2008).

211. Nuttall, G. A. & Houle, T. T. Liars, damn liars, and propensity scores. *Anesthesiology* **108**, 3–2 (2008).

212. D'Agostino, R. B. & D'Agostino, R. B. Estimating treatment effects using observational data. *J. Am. Med. Assoc.* **297**, 314 (2007).

213. Kappos, L. *et al.* A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *N. Engl. J. Med.* **362**, 387–401 (2010).

214. Gold, R. *et al.* Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis. *N. Engl. J. Med.* **367**, 1098–107 (2012).

215. Coles, A. J. Alemtuzumab Therapy for Multiple Sclerosis. *Neurotherapeutics* **10**, 29–33 (2013).

216. Marsh, E. A. *et al.* Alemtuzumab in the treatment of IVIG-dependent chronic inflammatory demyelinating polyneuropathy. *J. Neurol.* **257**, 913–919 (2010).

217. Schnitzer, T. J. *et al.* Subcutaneous administration of CAMPATH-1H: clinical and biological outcomes. *J. Rheumatol.* **24**, 1031–1036 (1997).

218. Lockwood, C. M. *et al.* Treatment of refractory Wegener's granulomatosis with humanized monoclonal antibodies. *QJM* **89**, 903–912 (1996).

219. Cox, A. L. *et al.* Lymphocyte homeostasis following therapeutic lymphocyte depletion in multiple sclerosis. *Eur. J. Immunol.* **35**, 3332–3342 (2005).

220. Hirst, C. L. *et al*. Campath 1-H treatment in patients with aggressive relapsing remitting multiple sclerosis. *J. Neurol.* **255**, 231–238 (2008).

221. Fox, E. J. *et al*. A single-arm, open-label study of alemtuzumab in treatment-refractory patients with multiple sclerosis. *Eur. J. Neurol.* **19**, 307–311 (2012).

222. Coles, A. J. *et al*. Alemtuzumab vs. interferon beta-1a in early multiple sclerosis. *N. Engl. J. Med.* **359**, 1786–1801 (2008).

223. Coles, A. J. *et al*. Alemtuzumab more effective than interferon ??-1a at 5-year follow-up of CAMMS223 Clinical Trial. *Neurology* **78**, 1069–1078 (2012).

224. Lycke, J. *et al*. Adverse event profile of alemtuzumab in active relapsing remitting multiple sclerosis patients who participated in the CARE-MS studies: Three-year follow-up. *Mult. Scler.* **19**, 487–488 (2013).

225. Brown, J. W. L. & Coles, A. J. Alemtuzumab: evidence for its potential in relapsing-remitting multiple sclerosis. *Drug Des. Devel. Ther.* **7**, 131–8 (2013).

226. Genzyme. *Healthcare Professional Guide: Using LEMTRADA (alemtuzumab) in patients with relapsing-remitting multiple sclerosis (RRMS)*. (2013).

227. McCormack, P. L. Natalizumab: A review of its use in the management of relapsing-remitting multiple sclerosis. *Drugs* **73**, 1463–1481 (2013).

228. Rudick, R. A. *et al*. Natalizumab plus interferon beta-1a for relapsing multiple sclerosis. *N. Engl. J. Med.* **354**, 911–923 (2006).

229. Havrdova, E. *et al*. Effect of natalizumab on clinical and radiological disease activity in multiple sclerosis: a retrospective analysis of the Natalizumab Safety and Efficacy in Relapsing-Remitting Multiple Sclerosis (AFFIRM) study. *Lancet Neurol.* **8**, 254–260 (2009).

230. Hutchinson, M. *et al*. The efficacy of natalizumab in patients with relapsing multiple sclerosis: Subgroup analyses of AFFIRM and SENTINEL. *J. Neurol.* **256**, 405–415 (2009).

231. National Institute for Health and Clinical Excellence (NICE) Technology Appraisal 127. Natalizumab for the treatment of adults with highly active relapsing-remitting multiple sclerosis. (2007). Available at: http://www.nice.org.uk/nicemedia/live/11822/36136/36136.pdf.

232. Tan, C. S. & Koralnik, I. J. Progressive multifocal leukoencephalopathy and other disorders caused by JC virus: clinical features and pathogenesis. *Lancet Neurol.* **9**, 425–437 (2010).

233. Barts and The London School of Medicine and Dentistry. Multiple Sclerosis research: Natalizumab PML update. 2017. Available at: http://multiple-sclerosis-research.blogspot.com/2017/11/latest-natalizumab-pml-risk-update.html.

234. Isidoro, L., Pires, P., Rito, L. & Cordeiro, G. Progressive multifocal leukoencephalopathy in a patient with chronic lymphocytic leukaemia treated with alemtuzumab. *BMJ Case Rep.* **12**, 164–79 (2014).

235. Wingerchuk, D. M. & Carter, J. L. Multiple sclerosis: current and emerging disease-modifying therapies and treatment strategies. *Mayo Clinic proceedings* **89**, 225–240 (2014).

236. Liu, C., Li Wan Po, A. & Blumhardt, L. D. 'Summary measure' statistic for assessing the outcome of treatment trials in relapsing-remitting multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **64**, 726–729 (1998).

237. Tuohy, O., Costelloe, L., Hill-Cawthorne, G. & Bjornson, I. Alemtuzumab treatment of multiple sclerosis: long-term safety and efficacy. *J. Neurol. Neurosurg. Psychiatry* **Published**, (2014).

238. Harris, J. O., Frank, J. A., Patronas, N., McFarlin, D. E. & McFarland, H. F. Serial gadolinium-enhanced magnetic resonance imaging scans in patients with early, relapsing-remitting multiple sclerosis: Implications for clinical trials and natural history. *Ann. Neurol.* **29**, 548–555 (1991).

239. Kappos, L. *et al.* Switching from natalizumab to fingolimod : A randomized, placebo-controlled study in RRMS. *Neurology* **85**, 29–39 (2015).

240. Kappos, L. *et al*. Factors influencing long-term outcomes in relapsing-remitting multiple sclerosis: PRISMS-15. *J. Neurol. Neurosurg. Psychiatry* **86**, 1202–1207 (2015).

241. European Medicines Agency. *Measures to minimise risk of serious side effects of multiple sclerosis medicine Lemtrada*. doi:EMA/609015/2019

242. La Ganke, C. & Adcock, A. Clinical outcomes of 200 multiple sclerosis patients switching from natalizumab to alemtuzumab in a single United States MS center. in *AAN* 86 (16 SUPPL 1) (2016).

243. Berger, T. *et al*. Alemtuzumab Use in Clinical Practice: Recommendations from European Multiple Sclerosis Experts. *CNS Drugs* 1–18 (2016). doi:10.1007/s40263-016-0394-8

244. Kelly, SB Kinsella, K, Duggan, M, Hutchinson M, Turbidy N, M. C. Initial experience of alemtuzumab for patients who fail natalizumab therapy [Abstract]. *J. Neurol. Neurosurg. Psychiatry* **83**, e1 (2012).

245. Fortin, J., Jaques, F., Sikati Fokko, V. & Grand'Maison, F. Switching from Natalizumab to Alemtuzumab in patinets with RRMS: Real-world experience. in *ECTRIMS* 22 (pp 783) (2016).

246. Malucchi, S. *et al*. High-Risk PML Patients Switching from Natalizumab to Alemtuzumab: an Observational Study. *Neurol. Ther.* **6**, 145–152 (2017).

247. Sorensen, P. S. *et al*. Recurrence or rebound of clinical relapses after discontinuation of natalizumab therapy in highly active MS patients. *J. Neurol.* **261**, 1170–1177 (2014).

248. Giovannoni, G. *et al*. Switching patients at high risk of PML from natalizumab to another disease-modifying therapy. *Pract. Neurol.* **16**, 389–393 (2016).