IMI WP5 Report 1:b:iv Benefit-Risk Wave 1 case study report: NATALIZUMAB

Risk benefit case study with a focus on testing methodology.

Four of four: natalizumab

Authors

Richard Nixon, Thai Son Tong Nguyen (Novartis)
Isabelle Stoeckert, Christoph Dierig, Silvia Kuhls (Bayer)
Gemma Hodgson (Takeda)
John Pears (Astra Zeneca)
Ioanna Tzoulaki (Imperial)
Dolores Montero (Agencia Española de Medicamentos y Productos Sanitarios)

On behalf of PROTECT Work Package 5 participants

Version 1 Date: 01 Mar 2013 Date of any subsequent amendments below	Person making amendments	Brief description of amendments

Acknowledgements: The research leading to these results was conducted as part of the PROTECT consortium (Pharmacoepidemiological Research on Outcomes of Therapeutics by a European ConsorTium, www.imi-protect.eu) which is a public-private partnership coordinated by the European Medicines Agency.
The PROTECT project has received support from the Innovative Medicines Initiative Joint Undertaking (www.imi.europa.eu) under Grant Agreement n° 115004, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution

Contents

Glossary 4
Section 1 Introduction and background 6
Part 1.1 Organization of Report 6
Part 1.2 Multiple Sclerosis Disease Background 6
1.2.1 Concept and Pathology 6
1.2.2 Diagnosis and Forms of Multiple Sclerosis 6
1.2.3 Medical need and available treatment options 7
1.2.4 Natalizumab and Progressive multifocal leukoencephalopathy 7
Section 2 Aim and objectives 11
Part 2.1 Key questions to be addressed 11
Section 3 Methods 12
Part 3.1 Justifications for selection of benefit-risk approaches 12
Part 3.2 Overview and analysis approach 13
Section 4 Evidence data 15
Part 4.1 BRAT Step 1 (Define the decision context) and PrOACT-URL (Step 1 and 3 Problem and Alternatives) 15
Part 4.2 BRAT Step 2 (Identify benefit and risk outcomes). PrOACT-URL Step 2 (Objectives) 16
Part 4.3 Gathering objective data: BRAT Step 3 (Identify and extract source data), PrOACT Step 4 (Consequences) 16
4.3.1 Literature searching 17
Part 4.4 Gathering subjective data: BRAT step 5 (assess outcome importance) and PrOACT step 5 (Trade offs). 23
4.4.1 Value functions 23
4.4.2 Weights 24
Section 5 Results 26
Part 5.1 BRAT framework using NNT/NNH 26
5.1.1 Introduction 26
5.1.2 BRAT Step 1: Define the decision context 27
5.1.3 BRAT Step 2 Identify benefit and risk outcomes 27
5.1.4 BRAT Step 3 Identify and extract source data 27
5.1.5 BRAT Step 4 Customize framework 29
5.1.6 BRAT Step 5: Assess outcome importance 29
5.1.7 BRAT Step 6: DISPLAY AND INTERPRET KEY BENEFIT-RISK MEASURES 29
5.1.8 Conduction of benefit-risk analysis based on the NNT/NNH concept 35

Part 5.2 PrOACT-URL framework using MCDA 38
5.2.1 Introduction 38
5.2.2 Step 1: Problem 39
5.2.3 Step 2: Objectives 39
5.2.4 Step 3: Alternatives 39
5.2.5 Step 4: Consequences 39
5.2.6 Step 5: Trade offs 39
5.2.7 Step 6: Uncertainty 43
5.2.8 Step 7: Risk tolerance, Step 8: Linked decisions 46
Section 6 Discussion 47
Part 6.1 Methodology: Appropriate frame 47
6.1.1 Meaningful reliable information 51
6.1.2 Clear values and trade-offs 54
6.1.3 Logically correct reasoning 60
6.1.4 Commitment to action 62
Part 6.2 The assessment of benefit-risk balance 63
6.2.1 BRAT Framework using weighted Net Clinical Benefit. 63
6.2.2 PrOACT-URL framework using MCDA. 63
Part 6.3 Visual representation of benefit-risk assessment results 63
6.3.1 BRAT 63
6.3.2 PrOACT-URL 63
Section 7 Conclusion 65
Section 8 References 67
Section 9 Appendix 68
Part 9.1 Data extraction 68
Part 9.2 Example questionnaire 73
Part 9.3 Weights elicited 74
Part 9.4 People 76

Glossary

AE	Adverse Effect
BRA	Benefit-Risk Analysis
BRAT	Benefit-Risk Action Team
CHMP	Committee for Medicinal Products for Human Use
CIS	Clinically Isolated Syndrome
CNS	Central Nervous System
DALY	Disability Adjusted Life Year
DMD	Disease-Modifyng Drug
DSA	Deterministic Sensitivity Analysis
EDSS	Expanded Disability Status Scale
EMA	European Medicines Agency
EMSP	European Multiple-Sclerosis Platform
EPAR	European Public Assessment Report
EU	European
FDA	Food and Drug Administration
GA	Glatiramer Acetate
Gd	Gadolinium
HR	Hazard Ratio
IFN	Interferon
IMI	Innovative Medicines Initiative
ITT	Intention to Treat
LCI	Lower Confidence Interval
MCDA	Multiple Criteria Decision Analysis
MRI	Magnetic Resonance Imaging
MS	Multiple Sclerosis
MSFC	Multiple Sclerosis Functional Composite
NCB	Net Clinical Benefit
NNH	Number Needed To Harm

NNT	Number Needed To Treat
Pbo	Placebo
PD	Pharmacodynamic
PK	Pharmacokinetic
PML	Progressive Multifocal Leukoencephalopathy
PPMS	Primary Progressive Multiple Sclerosis
PSA	Probabilistic Sensitivity Analysis
QALY	Quality Adjusted Life Year
RR	Relative Risk
RRMS	Relapsing-Remitting Multiple Sclerosis
SPMS	Secondary Progressive Multiple Sclerosis
Tx	Active treatment
UCI	Upper Confidence Interval
UK	United Kingdom
US	United States
VAS	Visual Analogue Scale
w-NCB	Weighted Net Clinical Benefit

Section 1 Introduction and background

Part 1.1 Organization of Report

This report details a case study for assessing the benefit-risk of natalizumab for the treatment of relapsing remitting multiple sclerosis. In section 1 the disease is described and the background and issues of the key safety concerns of Progressive Multifocal leukoencephalopathy described. Section 2 details the aims of the case study and the key questions to be addressed. Section 3 gives an overview and justification of the methods used. Sections 4 describe the raw data, both objective and subjective, that is used in the case study. Section 5 describes the results from applying two different methods of assessing the benefit-risk of natalizumab. Section 6 critiques the two methods with section 7 giving overall conclusions.

This purpose of this report is to assess methodology for benefit-risk assessment. Natalizumab is used as the issue of a rare serious side effect in an effective treatment for a serious disease makes for an interesting case study. The purpose of this report is not to revisit decisions made by regulatory authorities.

Part 1.2 Multiple Sclerosis Disease Background

1.2.1 Concept and Pathology

Multiple Sclerosis (MS) is the most common neurological disorder causing disability in young adults affecting approx. 1 in 1.000 people in western countries. MS begins at the age of 20 to 40 years with a median age of 28 years at onset with acute episodes of neurological dysfunction, followed by periods of partial or complete remission and clinical stability in between relapses (RRMS) (1). MS is an autoimmune disease that occurs in genetically susceptible individuals exposed to specific environmental triggers. Apart from demyelination by autoimmune disturbance, the other pathologic hallmark of the disease is inflammation. According to a strictly immunological explanation of MS, the inflammatory process is caused by T cells. Those trigger inflammatory processes, stimulating other immune cells and soluble factors like cytokines and antibodies. Leaks form in the blood–brain barrier, which in turn cause a number of other damaging effects such as swelling, activation of macrophages, and more activation of cytokines and other destructive proteins. Approximately 85% of patients suffer from relapsing-remitting MS. However, this descriptive term is misleading because ultimately signs and symptoms of MS tend to progress, highlighting the need for prompt and consistent therapy. Early treatment has been advocated, including for clinically isolated syndrome (CIS), often a precursor to MS (2).

1.2.2 Diagnosis and Forms of Multiple Sclerosis

A guideline from the regulators perspective provides a clear definition of the indication under discussion (3). The term relapsing MS includes 1) patients with RRMS, 2) patients with Secondary progressive multiple sclerosis (SPMS) and superimposed relapses and 3) patients with a single demyelinating clinical event who show lesion dissemination on subsequent MRI scans according to McDonald's criteria. Prevention and/or modification of relapse features as well as prevention or delay of the accumulation of disability as sequela of acute relapses, are meaningful goals in the treatment of RMS. Progression of disability, as a result of relapses from which patients do not fully recover take many years and, for the moment, there are no surrogate variables for evaluating progression of disability.

Therefore large-scale long-term parallel group trials are required to establish clinically relevant treatment differences on disease progression. Relapse rate, relapse duration, and recovery after relapses are all highly variable between patients and for individual patients. Therefore, treatments intended to decrease the relapse rate or modify relapses should be evaluated in parallel trials sufficiently large and long to overcome this inter and intra-individual variability.

Currently approved therapies have demonstrated a favourable effect on the rate and severity of relapses with some products similar in the short-term (a few years) progression of disability. If a product demonstrates a benefit in relapse rate or severity without an accompanying effect on preventing or delaying disability, the clinical relevance of such benefit should be justified. It is therefore accepted that the indication in relapsing MS will mainly rely on the effects shown in patients with RRMS, but may be extrapolated to an effect on relapses in SPMS.

Several major placebo-controlled clinical trials have provided evidence of an apparent short-term stabilisation in placebo-treated patients that could be explained, among others, by the regression to the mean phenomenon, and by a real placebo effect, as well as by the natural course of the disease. Approved therapies have been shown to favourably modify the short-term evolution of the disease although the benefit is modest, at the cost of significant patient inconveniences and side effects. Differences from placebo seem not always consistent across trials and the sensitivity of the available scales to measure progression of disability as well as other characteristics of clinical trials in this field do not assure the ability to detect clinically relevant differences.

1.2.3 Medical need and available treatment options

In MS, disability accrues during both acute inflammatory events and progressive phases of disease. In short to intermediate-term studies, suppression of acute inflammatory events reduced the risk of disability progression in MS. Therefore, early recognition of sub-optimal response to disease-modifying therapy (DMT) and prompt intervention may limit future impairment. Interventions for the treatment of MS include immunomodulatory agents such as interferon-beta or glatiramer acetate, monoclonal antibodies such as natalizumab or immunsuppressants such as mitoxantrone, azathioprine or cyclophosphamide.Very recently, fingolimod was approved and added to the list of disease-modifying agents. All of these agents reduce the frequency and severity of relapses to varying degrees, delay disability progression and decrease brain lesion development as evidenced by MRI. The Multiple Sclerosis Consensus Therapy Group (MSTCG) scheme considers interferon beta and glatiramer acetate as basic therapies whereas natalizumab and mitoxantrone represent possibilities for escalating treatment.

Natalizumab (Tysabri®) is a monoclonal antibody against α4-integrin, a component of VLA-4 located on leukocytes. Binding to VLA-4 blocks its interaction with the ligand VCAM (vascular cell adhesion molecule) on the surface of endothelial cells at the blood-brain barrier, thus greatly reducing the transmigration of lymphocytes and monocytes from venules and capillaries into inflamed tissue. In EU, natalizumab is indicated as a “disease-modifying monotherapy of highly active relapsing MS” for the following patient groups: 1) patients showing high levels of disease activity despite treatment with an IFN-β preparation, or 2) untreated/treatment-naive patients with rapidly progressing relapsing-remitting MS (at least two serious relapses per year).

1.2.4 Natalizumab and Progressive multifocal leukoencephalopathy

Progressive multifocal leukoencephalopathy (PML) is a brain infection caused by activation of the JC Virus (JCV), which is a type of human polyomavirus that is commonly found in the general population, but only leads to PML if the immune system has been compromised. Progressive multifocal leukoencephalopathy is a clinical manifestation of direct and active JCV infection of the oligodendrocytes, which leads to decreased myelin production and subsequent demyelination with resulting severe disability and ultimately, death. The signs and symptoms of PML can initially be clinically very similar to those of an MS attack due to the demyelination.

In October 2009 the CHMP together with the European Commission initiated an Article 20 review of the benefits and risks for natalizumab in view of the cases of PML that have been observed since natalizumab has been on the market and in consideration of the occurrence of Immune Reconstitution Inflammatory Syndrome (IRIS) in these patients once natalizumab has been stopped and Plasma Exchange (PLEX) and/or immunoabsorption (IA) was implemented.

The discussion on the validity as confirmatory or not of a positive CSF PCR for JCV virus was also discussed at the CHMP. The conclusion was that a negative CSF PCR for JCV virus, due to a potentially low number of viral copies, does not automatically mean the exclusion of PML.

1.2.4.1 PML data and discussion at time of initial EU approval

At the time of initial approval in 2006, two cases of PML, one fatal, had been reported by the Applicant in two patients with RRMS, both treated with a combination of natalizumab and beta-interferon for more than 2 years. A third case of PML was later discovered upon re-evaluation of the Crohn's Disease safety database in a subject that was originally presumed to have died of a malignant astrocytoma. No confirmed cases of PML in the mono-therapy MS clinical trials were identified in the post-marketing setting. Over 4,500 person-years of experience with natalizumab exist between the MS and Crohn's Disease programmes. Of the 1,617 MS subjects treated with natalizumab in the placebo-controlled experience, the majority (1,271 or 79%) have received the fixed dose of 300 mg; 1,123 (69%) have been followed for at least 1 year, and 1,062 (66%) have been followed for over 2 years, satisfying, and indeed exceeding ICH safety database requirements [EMA Scientific discussion 2006].

Based on the clinical trial data, the risk of PML at that time was calculated as 1 per 1.000 after 17.9 month of treatment and appears to increase with the number of infusions. It should be noted however that at this point in time PML was not associated with monotherapy and could therefore be considered a potential risk rather than an identified risk.

Based on the CHMP review of data on quality, safety and efficacy, the CHMP considered by consensus that the risk-benefit balance of natalizumab indicated as single disease modifying therapy in highly active relapsing remitting multiple sclerosis for the following patient groups:

- Patients with high disease activity despite treatment with a beta-interferon (SPC: see 5.1);

or

- Patients with rapidly evolving severe relapsing remitting multiple sclerosis (SPC: see 5.1).

was favourable and therefore recommended the granting of the marketing authorisation.

A subgroup analysis in AFFIRM Major Efficacy Endpoints (Patients with ≥2 relapses and ≥1 Gd-enhancing lesion) in 148 natalizumab treated patients versus 61 placebo treated patents confirmed this conclusion.

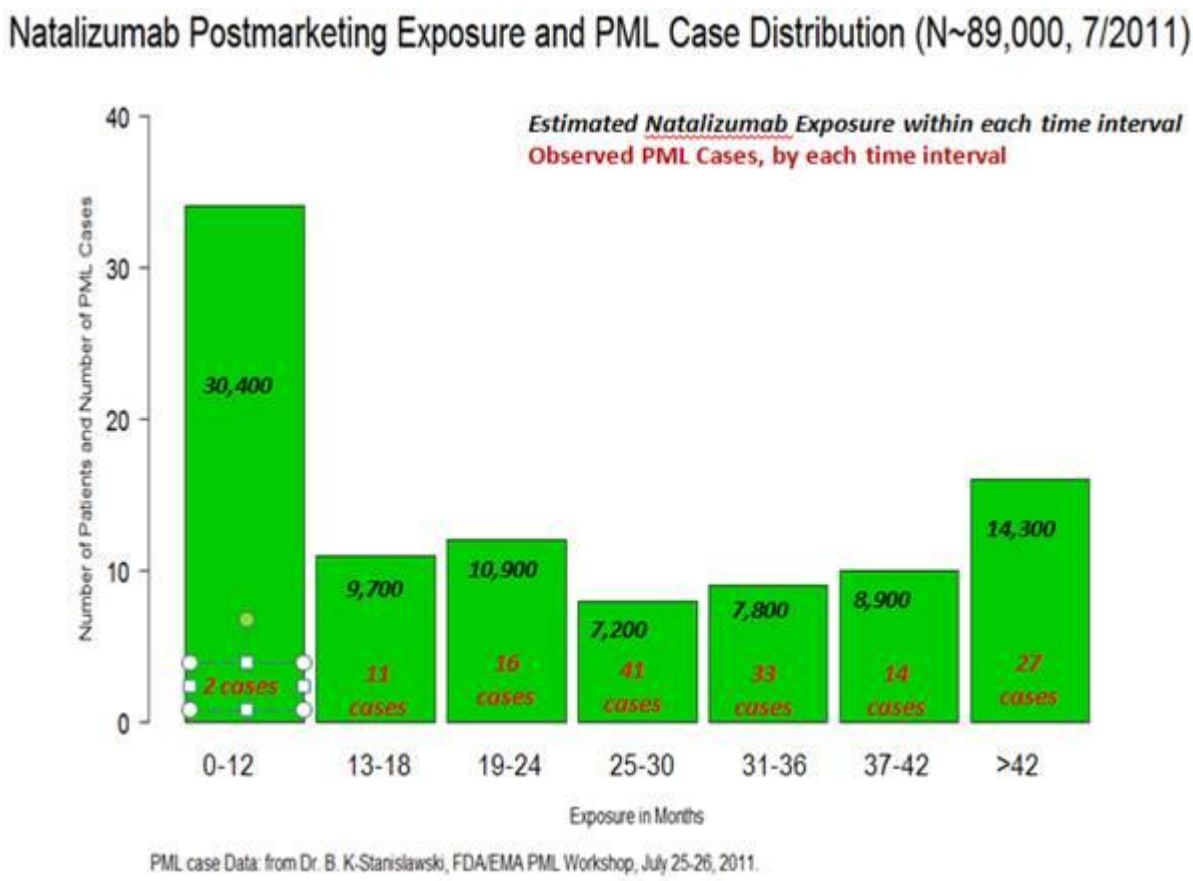
1.2.4.2 PML data at time of re-evaluation

Twenty-three confirmed cases of PML had been reported worldwide in patients with Multiple sclerosis (MS) receiving natalizumab between July 2008 and October 2009, resulting in four deaths. With increasing post-marketing experience and duration of exposure to natalizumab, the continued reporting of MS patients diagnosed with PML raised concerns, especially because these data suggest, that the risk for developing PML increases significantly after two years of continuous exposure.

The CHMP reviewed natalizumab after it had received reports of side effects in patients receiving the medicine. As mentioned above, these included 23 confirmed cases of PML reported worldwide between July 2008 and October 2009, resulting in four deaths. 14 of these cases, including one death, were reported in the EU. By the end of the review procedure on 20 January 2010, the total number of confirmed PML cases had risen to 31 worldwide, of whom 23 had been receiving natalizumab for more than two years. This is equivalent to around one case of PML for every 1,000 patients treated with natalizumab for two years or more. It has been seen that PML case frequency increases over time. At present it is observed that the frequency of PML cases increases dramatically after 24 months (monthly infusions) of treatment (4).:

As of May 4 2011 on the basis of post marketing reports, the estimated overall risk of PML has been reported as 1.51 per 1000 patients (95% CI 1.27-1.79) which is generally similar to rates seen in clinical trials (5).

More detailed observations were presented at a joint FDA / EMA workshop in July 2011. The graph below depicts in black, the number of patients exposed within each time interval, and in the red, the number of cases of PML observed with natalizumab, as of July 2011 (6).



1.2.4.3 CHMP conclusion

Having considered the overall data provided by the MAH in writing and in the oral explanation, the CHMP concluded that the benefit still outweighs the risks for the patients treated with natalizumab. The CHMP also concluded that the Product Information for natalizumab should include safety information aiming at informing patients and physicians about the risk of PML so that the symptoms are detected as soon as possible and therefore recommended the amendments to the relevant sections of the Summaries of Product Characteristics and Package Leaflet.

The Agency's Committee for Medicinal Products for Human Use (CHMP) has concluded that the risk of developing PML increases after two years of use of natalizumab, although this risk remains low. However, the benefits of the medicine continue to outweigh its risks for patients with highly active relapsing-remitting multiple sclerosis, for whom there are few treatment options available.

Because it is important that PML is detected early, the Committee recommended that a number of measures be put in place to ensure that patients and doctors are fully aware of the risks of PML (4).

1.2.4.4 *Patient involvement in decision making*

Public domain information from the website of the Europeans Medicines Agency indicates that: "The CHMP requested the views of patients and consumers on medicinal products under evaluation. This occurred three times in 2009; for the medicinal products Onsenal, Prezista and natalizumab. In each case, selected patients' organisations, fulfilling the Agency's criteria for interaction, were asked to answer in writing a list of questions adopted by the Committee, and in some cases were subsequently invited to participate in the CHMP discussion on the issues. These initiatives gathered the experience and views of the organisations on certain aspects of the current use of medicines, and this information has been taken into account by the CHMP on the opinions subsequently adopted and has been reflected in the CHMP assessment report "(7). It also indicates that natalizumab was an "example of patient participation in SAGs meeting, where very positive feedback on the patient contribution was collected (4).

Section 2 Aim and objectives

This document details a benefit-risk case study for natalizumab as part of the IMI PROTECT Work Package 5. natalizumab (natalizumab) was approved in 2004 by the FDA for the treatment of relapsing remitting multiple sclerosis (RRMS). In 2005 the drug was suspended because of an associated incidence of progressive multifocal leukoencephalopathy (PML), a rare neurological disorder. In 2006 it was re-introduced due to patient demand, but with strict risk minimization measures. In 2011 the inclusion of anti-JC virus antibody status as a PML risk factor was included in the label of natalizumab. (5)

In the EU, natalizumab was granted a Marketing Authorisation on 27 June 2006 for use as single disease modifying therapy in highly active relapsing remitting multiple sclerosis (RRMS) in patients with high disease activity despite treatment with a beta-interferon or in patients with rapidly evolving severe RRMS. The Regulatory decision making process in the re-evaluation based on the PML signal is further outlined below.

Part 2.1 Key questions to be addressed

Two potentially interesting questions are

1. Should natalizumab be given marketing approval at the time of first registration?
2. Should natalizumab be kept on the market given that increased episodes of PML were observed?

Section 3 Methods

PROTECT Work Stream B has recommended 13 approaches to be tested in the first wave of case studies. In this case study, we have considered the possibility of applying these approaches but it is not possible to apply every approach due to resources constraints.

Part 3.1 Justifications for selection of benefit-risk approaches

13 options of different methods have been recommended for testing in WP5. Each case study working group in WP5 will choose methods to apply in the first wave as they see appropriate, and these will be compared with the choices of the other case study groups to ensure that there is good coverage of all the options.

These 13 options cover four aspects of benefit-risk assessment. A different option may need to be picked from each of the four aspects for an assessment to be performed. Although it may be possible to perform a benefit-risk analysis omitting some of these aspects.

Table 1 Benefit-risk approaches included for testing in natalizumab case study

Aspect	BRAT subteam	ProACT subteam	Justification
Descriptive guidelines	Benefit Risk Action Team (BRAT) framework	PrOACT-URL framework	PrOACT-URL qualitative framework as MCDA fits well with this framework. BRAT is also applied so the two frameworks can be compared.
Benefit-risk assessment frameworks	Benefit Risk Action Team (BRAT) framework	Multi-Criteria Decision Analysis (MCDA).	MCDA as this is an intricate method capable of integrating multiple benefit and risk-criteria. It is one of the more complex methods. BRAT is a framework supported by a set of guidelines and a tool that allows to structure data for decision making.
Metric indices	Number Needed to Treat (NNT) and Number Needed to Harm (NNH).	Benefit-Risk Ratio (BRR).	MCDA naturally leads to a benefit-risk ratio (or actually benefit-risk balance as a ratio is not generally used in MCDA). NNT and NNH is a popular method which is relatively simple to apply. It would be insightful to assess how it compares to more principled methods for assessing benefit-risk, in particular looking at the improvements made to this method in the recent past.
Estimation techniques	Mixed Treatment Comparison (MTC).	Mixed Treatment Comparison (MTC).	Use the same methods for both frameworks so data can be shared between them.
Utility survey techniques	Direct elicitation	Direct elicitation	Use the same methods for both frameworks so weights can be shared between them. Direct elicitation is used. This is not one of the 13 options, but MCDA and the

	version of NNT needs a value elicitation method and discrete choice experiments are not a viable option given the time constraints and ethical approval needs for wave 1

Table 2 Benefit-risk approaches excluded from testing in natalizumab case study

Aspect	Option	Justification
Descriptive guidelines	Both covered	NA
Benefit-risk assessment frameworks	Stochastic Multi-criteria Acceptability Analysis (SMAA).	This is arguably the most complex method so better considered in the second wave of case studies
	Impact numbers.	
	Quality Adjusted Life Years (QALY).	
	Quality adjusted Time Without Symptoms and Toxicity (Q-TWiST).	A metric pertinent to oncology treatments.
	Incremental Net Health Benefit (INHB).	
Estimation techniques	Probabilistic Simulation Method (SPM)	
Utility survey techniques	Discrete Choice Experiment (DCE)	Discrete choice experiments are considered the current gold standard for value elicitation. However, they require a large study to be done on patients, and time does not permit this to be done.

Part 3.2 Overview and analysis approach

The team agreed that a framework is a prerequisite for appropriate decision making. The two frameworks were chosen to evaluate them in parallel and draw conclusions on suitability and similarities. The steps involved in these two methods are given below, and the steps that are very similar to each other are given on the same row of the table.

Table 3 Comparison of the BRAT and PrOACT-URL frameworks

BRAT framework with subsequent NNT/NNH /net clinical benefit analysis	MCDA using the PrOACT-URL framework
1) Define the decision context	1) Problem
2) Identify benefit and risk outcomes	2) Objectives

1) Define the decision context	3) Alternatives
3) Identify and extract source data	4) Consequences
4) Customize the framework	
5) Assess outcome importance	5) Trade offs
6) Display and interpret key benefit-risk metrics	6) Uncertainty
6) Display and interpret key benefit-risk metrics	7) Risk tolerance
6) Display and interpret key benefit-risk metrics	8) Linked decisions

Section 4 Evidence data

In order to define the data sources, we first needed to define the decision scope and identify the value tree. The same scope and value tree are used for both frameworks; this enables the use of the same objectives and subjective data to be used in the analysis.

Part 4.1 BRAT Step 1 (Define the decision context) and PrOACT-URL (Step 1 and 3 Problem and Alternatives)

The group discussed and agreed a proposal on objective and assumptions of the assessment and the contextual basis for it as outlined below:

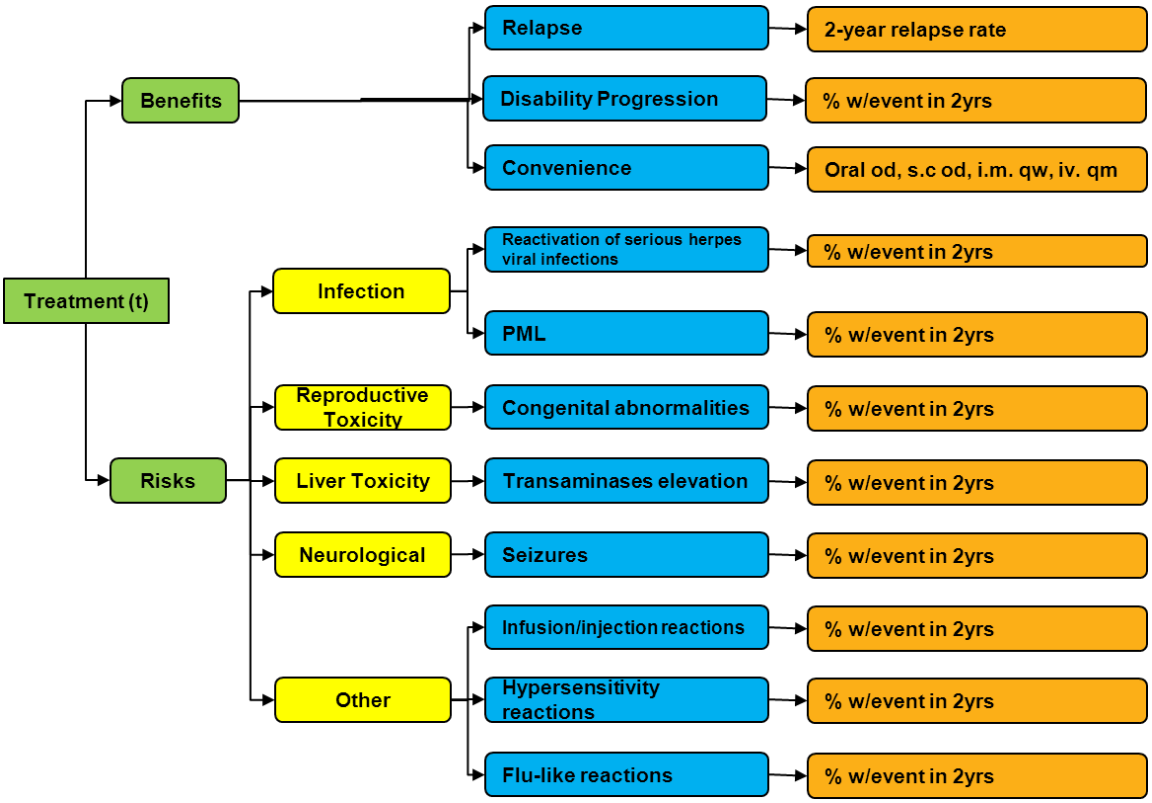
Table 4 Decision context resulting from the discussion

Objective	1. Should natalizumab be given marketing approval at the time of first registration? 2. Should natalizumab be kept on the market given that increased episodes of PML were observed?
Indication	Monotherapy for the treatment of highly active RRMS (relapsing remitting multiple sclerosis) SmPC section 4.1 indication in detail: Natalizumab is indicated as single disease modifying therapy in highly active relapsing remitting multiple sclerosis for the following patient groups: Adult patients aged 18 years and over with high disease activity despite treatment with a beta-interferon. These patients may be defined as those who have failed to respond to a full and adequate course (normally at least one year of treatment) of beta-interferon. Patients should have had at least 1 relapse in the previous year while on therapy, and have at least 9 T2-hyperintense lesions in cranial Magnetic Resonance Image (MRI) or at least 1 Gadolinium-enhancing lesion. A “non-responder” could also be defined as a patient with an unchanged or increased relapse rate or ongoing severe relapses, as compared to the previous year. or Adult patients aged 18 years and over with rapidly evolving severe relapsing remitting multiple sclerosis defined by 2 or more disabling relapses in one year, and with 1 or more Gadolinium enhancing lesions on brain MRI or a significant increase in T2 lesion load as compared to a previous recent MRI
Drug	Natalizumab
Formulation/Dose	Natalizumab 300mcg, iv, qm
Comparators	Interferon beta-1a: RMMS, 30mcg, im, qw Glatiramer acetate:RMMS, 20mg, sc, qd Placebo
Population	Patients with RRMS
Time Frame for Outcomes	24 months
Stakeholder	Decision maker Regulator (at EMA), taking the patient perspective. The regulator makes the decision, using the values and weights of a patient

Part 4.2 BRAT Step 2 (Identify benefit and risk outcomes). PrOACT-URL Step 2 (Objectives)

The value tree was prepared by some individual experts based on the available data. In general the Summary of Product Characteristics (SPC) for natalizumab and the European Assessment report by EMA were chosen as the foundation for this assessment as the data outlined could be considered as validated by Regulatory Authorities. A group decision was taken on the key benefits and risks to be included. Similarly it was agreed that PML should be included at time of simulated initial approval assessment as PML could be considered as a potential risk identified from combination treatment at this point in time.

Figure 1 Value tree



Part 4.3 Gathering objective data: BRAT Step 3 (Identify and extract source data), PrOACT Step 4 (Consequences)

This section details the sources of objective evidence this is used in the benefit-risk assessment. A two year time horizon for the decision was chosen as this is considered the minimum duration for the drugs in the scope to demonstrate efficacy. The scope is limited to Phase III studies as these have this time horizon and were thought to be able to provide the most reliable data and a sufficiently rich source of safety data for the analysis. The sources of evidence to extract relevant data from are:

- European public assessment reports (EPAR), “product information” and “scientific discussion”
- Literature search.

Periodic Safety Update Reports (PSURs) are not an option for wave 1 because of time restraints required to extract the relevant data.

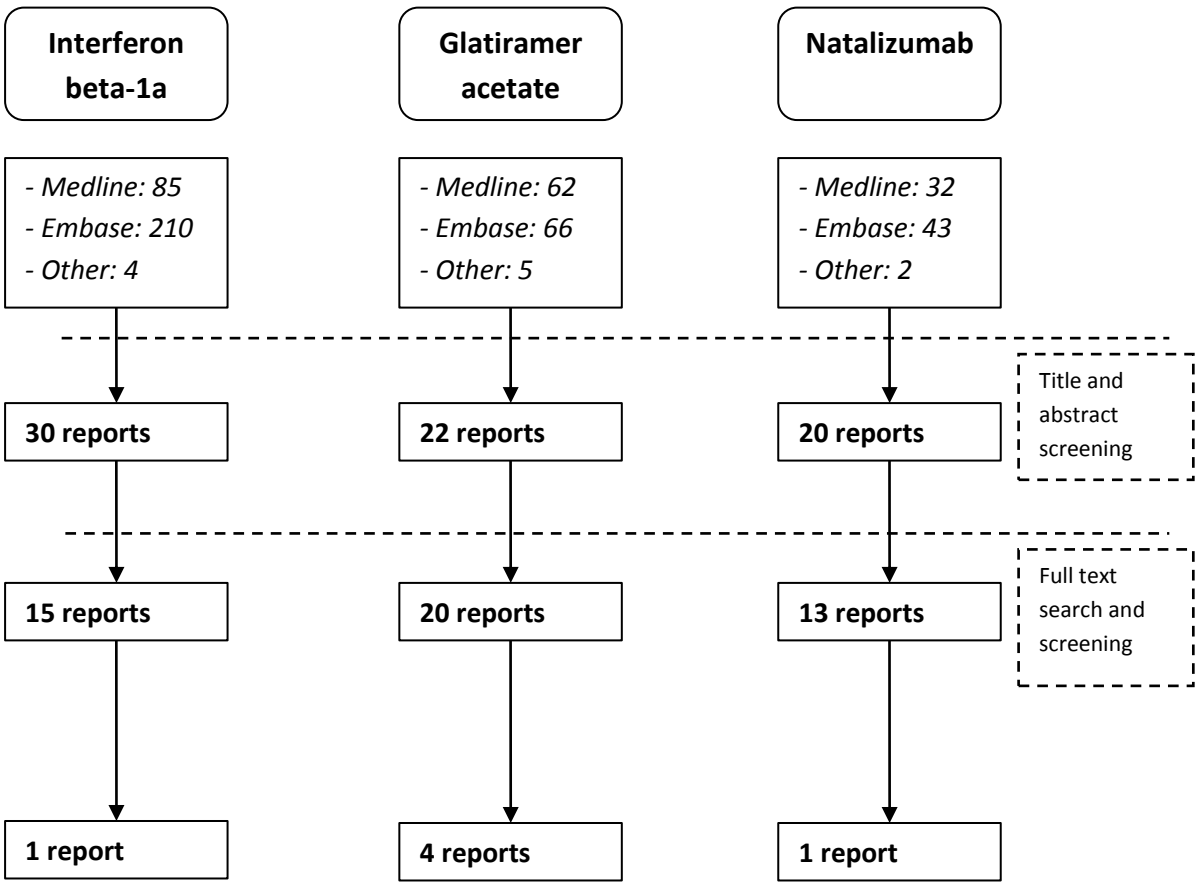
4.3.1 Literature searching

Two databases were searched from English-language reports of controlled clinical trials assessing the efficacy and safety of alternative MS drugs for relapsing-remitting multiple sclerosis: EMBASE from 1980 to July 2011 and Ovid MEDLINE from 1948 to July 2011. In addition, reference searching from review papers and expert's suggestion was also performed.

Database searching was performed for each drug using a similar search strategy, the details of these search terms is given in Table 21.

Results from literature searching and data collection for clinical trials are illustrated inFigure 2.

Figure 2 Literature searching results



At the final screening stage (full text screening), four reports of three Phase III placebo-controlled clinical trials have been found for glatiramer acetate while only one Phase III placebo-controlled trial has been found in each drugs for natalizumab and interferon beta-1a.

After discussion with clinicians and statisticians, three reports of glatiramer acetate were are excluded due to their methodology and quality. Rationale for trials exclusion at this stage is presented inTable 5.

Table 5 Excluded clinical trials

Trials	Rationale for exclusion
Bornstein 1987	Only pilot trial with small sample size and using match-pair randomization
Johnson 1998	Extension period over 2 years from the core trial (Johnson KP 1995)
Comi 2001	Trial period was only 9 months

Details of the studies from which data are extracted are given in Table 6

Table 6 Details of studies used in the analysis

Drug	Author	Year	Exposure	Comparator	Details
Natalizumab	Polman CH	2006	28 Months	Placebo	RRMS. Nov 2001: 942 patients - multiple centres and countries. 856 patients (91%) completed the 120-week trial
Interferon beta-1a	Jacobs LD	1996	24 Months	Placebo	RR MS. Early 1993: 301 patients - 4 clinical centers US.
Glatiramer acetate	Johnson KP	1995	24 months	Placebo	RRMS. October 1991-May 1992: 251 randomized. 215 completed 2-year trial. US population.

Efficacy data were only extracted from publications whereas safety data were also extracted from supporting data from FDA registration documents. The data source table is given in the Appendix in Table 22. This shows one row per outcome per study for each drug and the placebo. Cells shaded in yellow denote raw data extract from the source, in pink exact calculations performed on this data and cells not shaded denote that the number in them needed assumptions to be made in their calculations.

The data in the data source table is used to calculate the outcomes for each drug for each outcome on the scale of measurement given in the value tree. These data are given in the master data summary table in Table 7. It is necessary to measure the benefit and risk outcomes on an absolute scale, as this is the scale that values need to be elicited. However, the studies were performed in different patient populations and at different points in time, so comparing absolute effects between treatment arms directly will be confounding the effect of the drug and the effect of the patient population. It is common practice in biostatistics to assume that the relative treatment or safety effect is robust to changes in patient population. We compare drugs by taking a common placebo population, and finding the relevant absolute outcome given a drug by multiplying outcome in the placebo group with the appropriate relative rate. This is a simple form of network meta-analysis. The placebo group in the natalizumab trial is chosen as the common placebo group in most outcomes. However where there are missing data in natalizumab case, the common placebo group is chosen as the one with available data. For rare adverse events such as PML, herpes viral infections, or seizures, it has been assumed that no event has been observed in the common placebo group, and the unadjusted outcome in the treatment arm is used.

Table 7 Master data summary table

GROUP	CATEGORY	OUTCOME	MEASURE	DRUG	COMMON PLACEBO DRUG	COMMON PLACEBO EST	RELATIVE VALUE DESCRIPTION	RELATIVE VALUE EST	OUTCOME ON VALUE SCALE EST
Benefit	Relapse	Relapse	2 year relapse rate	Placebo	Natalizumab PBO	0.73[1]	rate ratio	1	1.46
				Natalizumab		0.73		0.32	0.47
				Interferon beta-1a		0.73		0.82	1.19
				Glatiramer acetate		0.73		0.71	1.04
	Disability progression	Disability progression	6-month confirmed proportion progressing after 2 years	Placebo	Natalizumab PBO	0.17	hazard ratio	1.00	0.23
				Natalizumab		0.17		0.46	0.11
				Interferon beta-1a		0.17		0.58	0.14
				Glatiramer acetate		0.17		0.77	0.18
	Convenience	Convenience	Route and	Placebo					oral od

			frequency of adminsitration	Natalizumab Interferon beta-1a Glatiramer acetate			iv qm hosp im qw sc od
Risk	Infection	Reactivation of serious herpes viral infections		Natalizumab PBO	event % 	0% 0% 0% 0%	event ratio 1 1.00 1.00	0% 0% 0% 0%
		PML						0% 0.151% 0% 0%
	Reproduction toxicity	Congenital abnormalities						0% 0.0% 0.0% 0.0%

	Liver Toxicity	Transaminases elevation	ALT >5x ULN	Natalizumab	event %	4%	event ratio	1	4.0%
				PBO		4%		1.25	5.0%
						4%		1.00	4.0%
						4%		1.00	4.0%
	Neurological	Seizures							0.5%
									0.5%
									3.0%
									0.0%
	Others	Infusion reactions/injection reactions		Natalizumab	event %	18%	event ratio	1	0.0%
				PBO		18%		1.34	23.6%
						18%		1.00	17.6%
						18%		1.53	26.9%
		Hypersensitivity Reactions		Natalizumab	event %	0%	event ratio	1.00	0.0%
				PBO		0%		0.00	0.0%
						0%		1.00	0.0%

					0%		1.50	0.0%
		Flu-like reactions		Interferon beta-1a PBO	event %	40%	event ratio	1.00	39.9%
						40%		1	39.9%
						40%		1.52	60.8%
						40%		1	39.9%

1) One year rate. Value tree outcome is the two year rate so this is doubled

Part 4.4 Gathering subjective data: BRAT step 5 (assess outcome importance) and PrOACT step 5 (Trade offs).

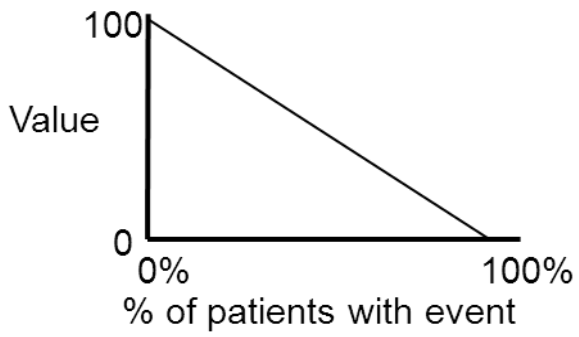
Ideally the relative importance of the benefit and risk outcomes would be elicited directly from patients. Discrete Choice Experiment is a currently recommended method for doing this. However, this requires enrolling many patients into a study, which precludes applying this method to this case study. An alternative approach is to elicit weights directly from patients, patient representatives or from experts with clinical knowledge who would take on the patient's perspective. As ethics approval is required from each country the patients are in and this would take more time and resources than we have for the case study, we elicited weights from patient representatives. This was performed at a decision conference held by teleconference on 23 September 2011 with two members of the European Multiple-Sclerosis Platform, facilitated by RN.

If the disparate measures used in the value tree are to be compared they need to be put on a common scale. This is performed in two steps. Firstly a value function is defined for each outcome which maps the outcome measure to a preference value scale. This defined the within-outcome importance for each outcome. Secondly weights are defined for each outcome, describing how important each outcome is relative to the other outcomes. This defines the between- outcome importance.

4.4.1 Value functions

For outcomes which have performance levels that can be measured by proportions (all outcomes except relapses and convenience), a linear value function is used. The range of possible outcome measure for proportions is 0 to 1. The best outcome is assigned a value of 100 (this is all patients experiencing the unfavourable event). The worst outcome is assigned a value of 0 (this is all patients avoiding the unfavourable event). This is implicitly assuming that each patient experiencing an unfavourable event is worth the same. An example of this value function is given in Figure 3. If the value function were convex, then the value of say 10 patients experiencing an event could be worth less than 10 times the value of one patient experiencing the event. However, for a regulator such a value function may be valid if, for example, there is an outcome with a threshold of incidence of adverse events, above which the drug is considered unacceptable. In this case the value function could have a value of zero for all proportions of patient that experience this event above this threshold.

Figure 3 Value function used for proportions



The 2-year relapse rate was also assigned a liner value function, but this had the best and worst outcomes defined at 0 and 2 relapses per two-years respectively.

The convenience outcome is on a categorical scale, and the value of each route and frequency of administrate is elicited at the same time as the weights. The elicited value function is given in Figure 4.

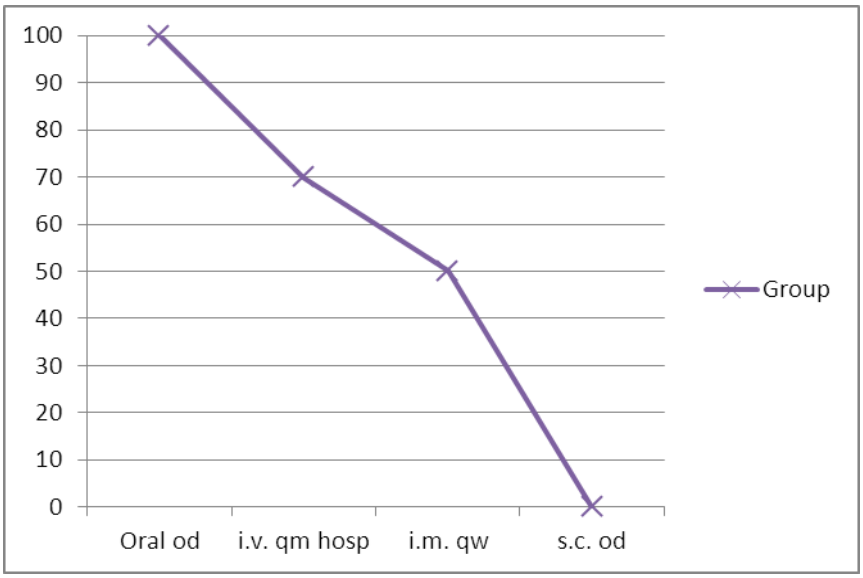


Figure 4 Value function for convenience

4.4.2 Weights

The weighting process was done using a bottom-up (or right-to-left approach given the orientation of our value tree) approach in the hierarchical value tree through decision conferences with the European Multiple-Sclerosis Platform (EMSP).

Firstly, the representatives of the EMSP were asked to take the regulator perspective making the decision in drug licensing on behalf of MS patients, however, the EMSP felt they could not take this perspective and instead gave their own values. The weighting process was organized in two meetings. The first meeting was to give an overview of the project and give instructions on completing the questionnaires. An example questionnaire is given in the Appendix in Part 9.2. Then the representatives completed the questionnaires individually, between the meetings. The second meeting was a group decision conference. The representatives presented their individual answers to the questionnaires and then came to a group consensus for the weights being elicited.

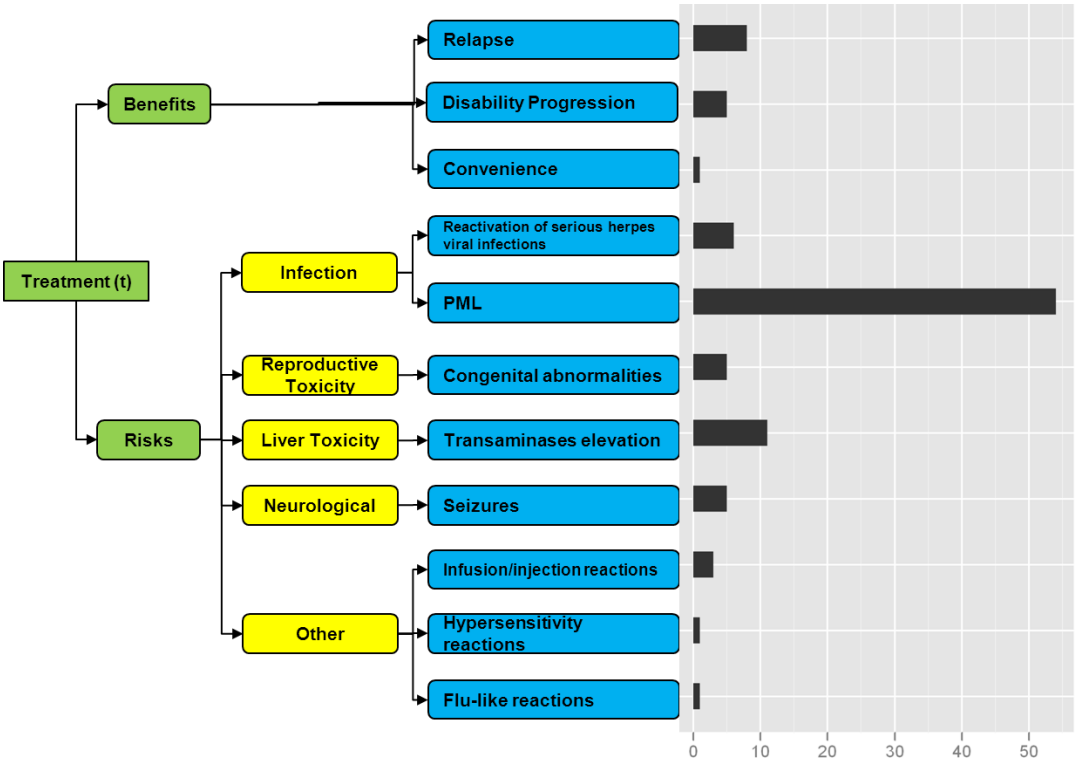
Weights are elicited thought the hierarchy of the value tree. Starting with each outcome measure, the relative ranks of a swing from the best to the worst of each outcome within each category are elected. Start from where all outcome scores are at the worst score, and choose the outcome you would most want to move to the best score. Then rank the other outcomes in a similar way. This is a “thought stepping stone” for then putting preference weights on these outcomes. For this, the top ranked outcomes is given a weight of 100, and place the other outcomes on the scale to reflect their relative importance.

The same approach is used to elicit weights between outcome categories, separately within benefits and risks. The top-ranked outcome from each category is used as a representative of its category. This top-ranked outcome in each category is ranked and weighted in the same way as before. Finally the benefits and risks are weighted using the same approach.

Cross checks were performed, where the weights were converted into the number of patients that would have to have a pair of event avoided to be in equipoise. For example “Imagine a clinical trial of 1000 patients with 1 patient developing PML in the treatment arm. How many patients would need to have an EDSS progression prevented for you to be indifferent about the benefit and harm caused by the treatment?”

The weights elicited at the different levels of the hierarchy are then used to calculate the overall cumulative weights for each outcome. This is shown in the Appendix in section **Error! Reference source not found.** Essentially the eight of an outcome is elicited, and then it is increased in proportion to how much weight it represented lower in the hierarchy. Note that the preference weight for relapses was elicited assuming a range out outcome between 0 and 1. As a range larger than this is needed, it is rescaled to the range 0 to 2, and so the swing weight needs to be doubled to account for this. The weights used for the benefit-risk analysis are shown in Figure 5. Note that as the weight for a relapse is for a value function with the measure scale with a range from 0 to 2, then actual weight of a single relapse is half that given in the Figure.

Figure 5 Preference weights for each outcome in the value tree



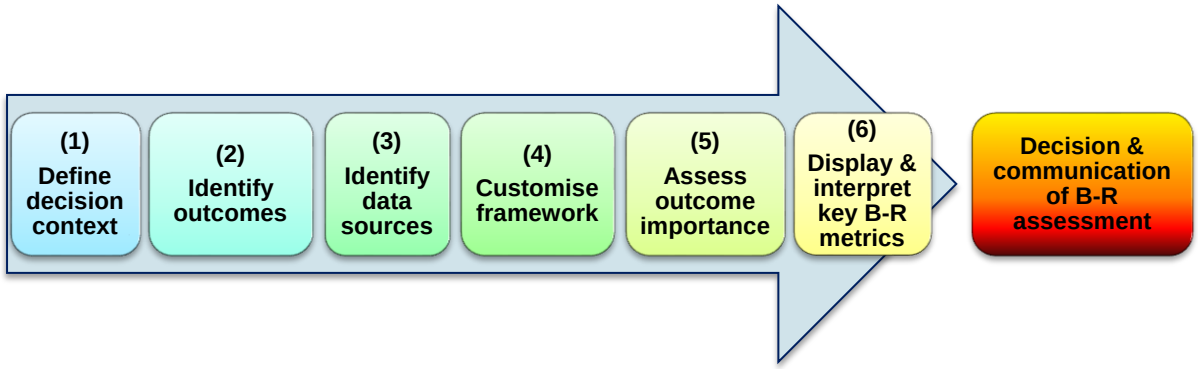
Section 5 Results

Part 5.1 BRAT framework using NNT/NNH

5.1.1 Introduction

The "BRAT" framework was developed by Pharmaceutical Research and Manufacturers of America (PhRMA) benefit-risk action team (BRAT) and aims to guide decision-makers in selecting, organizing, understanding and summarising the evidence relevant to benefit-risk decisions. BRAT can be considered a Benefit-risk assessment framework of mainly descriptive nature to conduct benefit-risk assessment in line with good decision-making practice in terms of preparation and transparency. It provides a guideline on organising, understanding and summarising evidence of benefits and risks into tabular outputs and graphical summaries to allow direct comparison and standardise the output for decision making. The guideline does not prescribe a particular setting for decision making itself and in particular proposes to avoid integration of benefits and risks as evidence in order to make it more accessible and transparent to those not familiar with complex statistical models. See also Coplan (2011) (8) as an introductory and core reference, and Levitan (2011) (9) for worked examples.

The key steps are illustrated below:.



The team followed the steps along the lines of the BRAT User's manual. In the following, each step was documented in terms of group results, summary of the team's discussion and comments/ assessment of the framework.

Table 8 Benefit Risk Action Team (BRAT) guidelines

Step	Description
1. Define the decision context	Define drug, dose, formulation, indication, patient population, comparator(s), time horizon for outcomes, perspective of the decision makers (regulator, sponsor, patient, or physician)
2. Identify outcomes	Select all important outcomes and create the initial value tree. Define a preliminary set of outcome measures/endpoints for each. Document rationale for outcomes included/excluded
3. Identify and extract source data	Determine and document all data sources (e.g. clinical trials, observational studies). Extract all relevant data for the data source table, including detailed references and any annotations, to help the subsequent interpretations

	create summary measures
4. Customise the framework	Modify the value tree on the basis of further review of the data and clinical expertise. Refine the outcome measures/endpoints. May include tuning of outcomes not considered relevant to a particular benefit-risk assessment or that vary in relevance by stakeholder group
5. Assess outcome importance	Apply or assess any ranking or weighting of outcome importance to decision makers or other stakeholders
6. Display and interpret key benefit-risk metrics	Summarise source data in tabular and graphical displays to aid review and interpretation. Challenge summary metrics, review source data, and identify and fill any information gaps. Interpret summary information.

5.1.2 BRAT Step 1: Define the decision context

See section Part 4.1 for details of the decision context and section Part 6.1 for a discussion of the process

5.1.3 BRAT Step 2 Identify benefit and risk outcomes

See section Part 4.2 for details of the value tree and section 6.1.1 for a discussion of the process.

The first step when entering the information into the BRAT tool is to create the value tree in the 'Value Tree'-sheet of the excel file. The tool only allows two main branches of the value tree, namely 'Benefits' and 'Risks'. Therefore 'Convenience' is entered as an outcome under 'Benefits' within the outcome category 'Convenience Benefits'. The benefits 'Relapse' and 'Disability Progression' are grouped in the category 'Medical Benefits'. Five risk categories are entered under 'Risks', where the categories 'Infection' and 'Other' have more than one associated outcome. The weights of the outcomes (see section 4.5) are entered after the outcome name.

5.1.4 BRAT Step 3 Identify and extract source data

See section Part 4.3 for details of the data and section 6.1.1 for a discussion of the process.

5.1.4.1 Entering data into BRAT Tool

After completion of the value tree as outlined above two filters are defined in the Excel Sheet 'Filters'. The first filter refers to the comparator and can take the values 'interferon beta-1a', 'glatiramer acetate' and 'placebo'. The second filter refers to the time point for the benefit risk assessment and can take the values 'At time of approval' and 'After approval'. By defining the filters it is possible to handle the data for all three comparisons of interest (natalizumab vs. interferon beta-1a, natalizumab vs. glatiramer acetate and natalizumab vs. placebo) and all time points of interest within the same file. In the 'Table settings' menu for the Benefit Risk summary table and the Forest plot the user can select which combination of filter values should be shown and then the respective data are displayed.

With the information entered in the 'Value Tree'- and 'Filters'- sheet the program automatically creates data rows in the 'Data'-Sheet that can be filled. Column names can be renamed by the user to clarify which estimates, confidence intervals, patient and event numbers refer to natalizumab and which refer to the comparator. Wherever possible the point estimates for natalizumab and the three comparators that were extracted and used in the PROACT URL work stream are entered into the tool. Please refer to Section 4.3.1 for details on how the data were extracted and transformed.

For the benefit '6-months sustained Disability Progression' and the risks 'Reactivation of serious herpes viral infections', 'PML', 'Transaminases elevation', 'Seizures', 'Infusion reactions/injection reactions', 'Hypersensitivity Reactions' and 'Flu-like reactions' the scale of the data provided by PORACT URL work stream is always 'proportion of subjects with the event in two years'. All data for these outcomes are entered into the tool without modification. As outlined earlier (see Section 4.3.1), indirect comparisons had to be done in order to allow comparisons of the active compounds as no head to head data are available.

The benefit 'Relapse' was given in a different scale as a '2 year relapse rate'. To get numbers in the same scale than the other outcomes, numbers for 'proportion of subjects with at least one relapse in two years' are extracted from the original publications. It would of course also be possible to directly use the '2 year relapse rate' in the quantitative benefit-risk assessment, however, this would imply that two relapses seen in one patient have the same weight in the assessment as two patients with one relapse each. Depending on the endpoint, this might be a reasonable assumption that should be scrutinized for the individual endpoint(s) under consideration.

Table 9 Proportion of subjects with at least one relapse in two years from original publications

Polman (2006)		Johnson (1995)		Jacobs (1996)	
Natalizumab	Placebo	Interferon beta-1a	Placebo	Glatiramer acetate	Placebo
0.28	0.54	0.62	0.74	0.664	0.73

After transformation (placebo adjustment, see also section 4.3.1) the numbers entered into the tool are shown in Table 4-4:

Table 10 Numbers for 'Relapse' after placebo adjustment

Treatment	Proportion of subjects with at least one relapse in two years
Natalizumab	0.28
Interferon beta-1a	0.45 (=0.54*(0.62/0.74))
Glatiramer acetate	0.49 (=0.54*(0.664/0.73))
Placebo	0.54

Two other outcomes given in a different scale are the categorical outcomes 'Congenital abnormalities' (with the potential values 'A', 'B', 'C', 'D' and 'X') and 'Convenience' (with the potential values 'iv qm', 'im qw' and 'sc qd'). The data fields for these outcomes are left empty in the BRAT tool as the outcomes cannot be transformed to the same scale as the other outcomes. Rows for 'Congenital abnormalities' and 'Convenience' appear in the Benefit Risk Summary table and Forest plot but without associated data.

For the risk 'PML' different data are entered for time point 'At time of approval' and 'After approval'. 'At time of approval' 3 PML events had occurred for natalizumab in 4500 patient years. As the number of patients is unknown we calculate the incidence proportion for natalizumab as the mean of 3/2250 and 3/4500 (i.e. assuming one time that 2250 patients were treated for two years and one time that 4500 patients were treated for one year) which results in a value of 0.001. As the events were not directly associated with monotherapy PML should be considered as potential risk at this stage. The incidence proportion for natalizumab for 'After approval' entered into the tool is

0.0015 and the risk can be considered an identified risk at that stage. For all other outcomes data for the two time points are the same.

Risk differences (point estimate natalizumab – point estimate comparator) are calculated for all outcomes where data are available. Where the number of patients and the number of events are available asymptotic confidence intervals for the risk difference are determined.

In those cases where the PROACT URL work stream assumed no difference between treatment and placebo the patient and event numbers for placebo were used in the confidence interval calculation (for example for outcome 'Reactivation of serious herpes viral infections' no difference between interferon beta-1a and Placebo from natalizumab study was assumed. Therefore the number of patients and number of events for Placebo from natalizumab study were used when calculating the confidence interval for the risk difference of natalizumab and interferon beta-1a). In some cases where confidence intervals are missing they might be obtained by simulation that takes into the account the transformation that was done in the placebo adjustment but such simulations are not within the timeframe of this case study and are therefore not done.

With the information entered in the 'Data'-sheet the BRAT tool automatically creates the Benefit Risk summary table and forest plot for the selected filter values (see Section 5.1.7). In the 'Global Settings'-sheet under 'Overall settings' the option 'Risk for proportions' has to be selected in the 'Type of data' field in order to change the column headers in the benefit risk summary table from 'per 1000 patient years' to 'per 1000 pts'.

See section 6.1.1 for a discussion of the process.

5.1.5 BRAT Step 4 Customize framework

5.1.5.1 Results

The Value Tree and the table are dynamic tools that can be revisited and changed as new information is acquired over the lifecycle of a compound. This includes, for example, such considerations as whether benefit and risk outcomes originally considered to be potential outcomes are either no longer included or become known outcomes. In this case, a filter was included in the source table to allow presentation of the data based on the lower PML occurrence assumed at time of approval and at the time of re-evaluation of benefit risk based on a higher PML occurrence. The adaptions of the framework can also be considered to adapt the visual output provided in tables and forest plot to support decision making.

See section 6.1.1 for a discussion of the process.

5.1.6 BRAT Step 5: Assess outcome importance

See section 4.4 for details of assessment of outcome importance and section 6.1.1 for a discussion of the process.

5.1.7 BRAT Step 6: DISPLAY AND INTERPRET KEY BENEFIT-RISK MEASURES

5.1.7.1 Results

Numerous approaches are available for summarizing and displaying information in benefit-risk assessments. The BRAT tool proposes a key summary table and a forest plot for visualization and decision making.

The table can be adapted as needed and was designed to provide an overview on all key available data at one glance. The example below describes natalizumab events / 1000 patients versus placebo events / 1000 patients at time of CHMP re-evaluation. Similar tables are available for comparison with other treatments (see Table 7.1-7.3. All tables include the weights provided by the patient association in the column to the right. Similar tables were also generated for time of approval as well (not displayed).

Table 11 Natalizumab versus placebo (Comparator) at time of CHMP re-evaluation

		Outcome	Natalizumab Risk / 1000 pts	Comparator Risk / 1000 pts	Risk Difference (95% CI)/ 1000 pts	
Benefits	Convenience Benefits	Convenience (weight 0.6%)	-	-	-	(-, -)
	Medical Benefits	Relapse (weight 3.9%)	280	540	-260	(-326, -195)
		Disability Progression (weight 5.6%)	110	230	-120	(-, -)
Risks	Infection	Reactivation of serious herpes viral infections (weight 6.7%)	80	70	10	(-26, 45)
		PML (weight 55.9%)	2	0	2	(-, -)
	Liver Toxicity	Transaminases elevation (weight 11.2%)	50	40	10	(-16, 38)
	Reproductive Toxicity	Congenital abnormalities (weight 5.6%)	-	-	-	(-, -)
	Neurological Disorders	Seizures (weight 5.6%)	0	0	0	(-, -)
	Other	Infusion/Injection reactions (weight 2.8%)	236	180	56	(6, 114)
		Hypersensitivity reactions (weight 1.1%)	90	40	50	(20, 82)
		Flu-like reactions (weight 1.1%)	399	400	-1	(-114, 114)

Higher for Natalizumab

Higher for Comparator

Table 12 Natalizumab versus glatiramer acetate (Comparator) at time of CHMP re-evaluation

		Outcome	Natalizumab Risk / 1000 pts	Comparator Risk / 1000 pts	Risk Difference (95% CI)/ 1000 pts	
Benefits	Convenience Benefits	Convenience (weight 0.6%)	-	-	-	(-, -)
	Medical Benefits	Relapse (weight 3.9%)	280	490	-210	(-, -)
		Disability Progression (weight 5.6%)	110	180	-70	(-, -)
Risks	Infection	Reactivation of serious herpes viral infections (weight 6.7%)	80	70	10	(-26, 45)
		PML (weight 55.9%)	2	0	2	(-, -)
	Liver Toxicity	Transaminases elevation (weight 11.2%)	50	40	10	(-16, 38)
	Reproductive Toxicity	Congenital abnormalities (weight 5.6%)	-	-	-	(-, -)
	Neurological Disorders	Seizures (weight 5.6%)	0	0	0	(-, -)
	Other	Infusion/Injection reactions (weight 2.8%)	236	269	-33	(-, -)
		Hypersensitivity reactions (weight 1.1%)	90	60	30	(-, -)
		Flu-like reactions (weight 1.1%)	399	399	0	(-114, 114)

Higher for Natalizumab
Higher for Comparator

Table 13 Natalizumab versus interferon beta-1a (Comparator) at time of CHMP re-evaluation

		Outcome	Natalizumab Risk / 1000 pts	Comparator Risk / 1000 pts	Risk Difference (95% CI)/ 1000 pts	
Benefits	Convenience Benefits	Convenience (weight 0.6%)	-	-	-	(-, -)
	Medical Benefits	Relapse (weight 3.9%)	280	450	-170	(-, -)
		Disability Progression (weight 5.6%)	110	140	-30	(-, -)
Risks	Infection	Reactivation of serious herpes viral infections (weight 6.7%)	80	70	10	(-26, 45)
		PML (weight 55.9%)	2	0	2	(-, -)
	Liver Toxicity	Transaminases elevation (weight 11.2%)	50	40	10	(-16, 38)
	Reproductive Toxicity	Congenital abnormalities (weight 5.6%)	-	-	-	(-, -)
	Neurological Disorders	Seizures (weight 5.6%)	0	11	-11	(-23, 0)
	Other	Infusion/Injection reactions (weight 2.8%)	236	312	-76	(-, -)
		Hypersensitivity reactions (weight 1.1%)	90	40	50	(20, 82)
		Flu-like reactions (weight 1.1%)	399	608	-209	(-320, -98)

Higher for Natalizumab
Higher for Comparator

It should be noted that in the following figures comparing active compounds, bars might be missing because indirect comparisons had to be performed and information on variability for such cases is missing. Such missing data are captured as warnings by the BRAT tool in a separate table to highlight some uncertainties.

Forest plot display - A forest plot can instantly give the viewer a sense of the magnitude of difference and the variability or precision of a measure.

Figure 6 Natalizumab versus placebo (Comparator) at time of CHMP re-evaluation

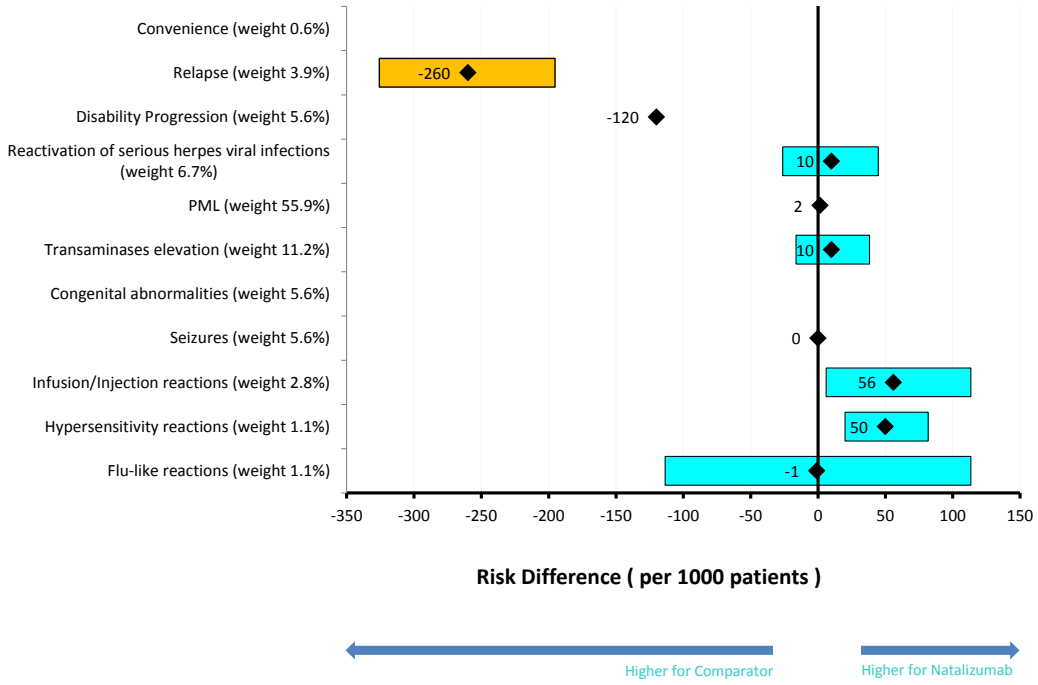


Figure 7: Natalizumab versus glatiramer acetate (Comparator) at time of CHMP re-evaluation

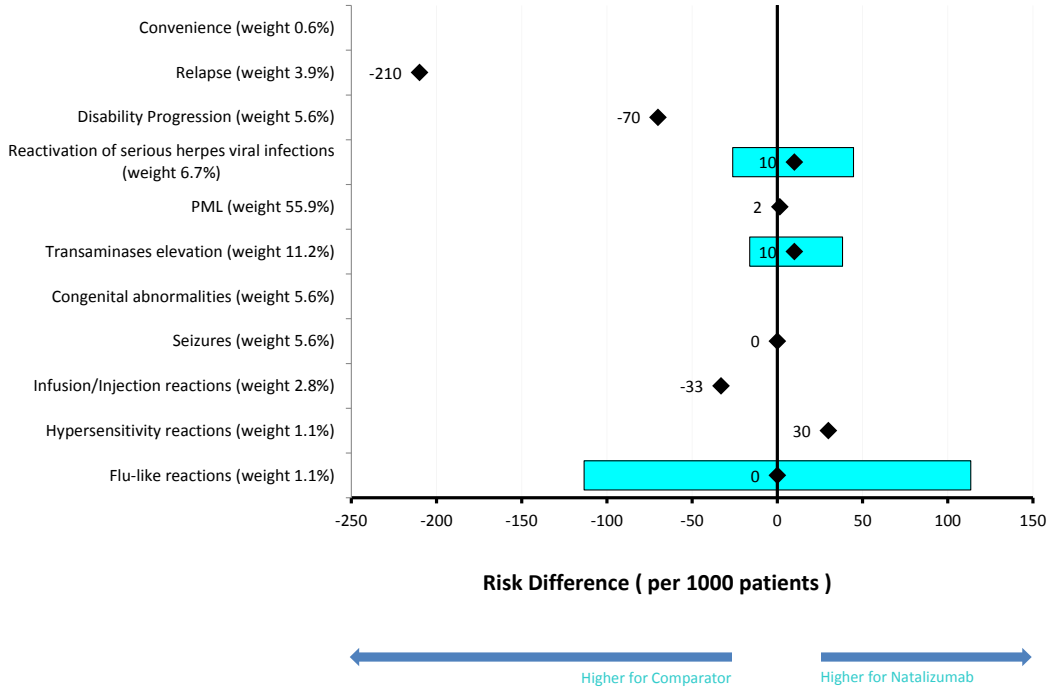
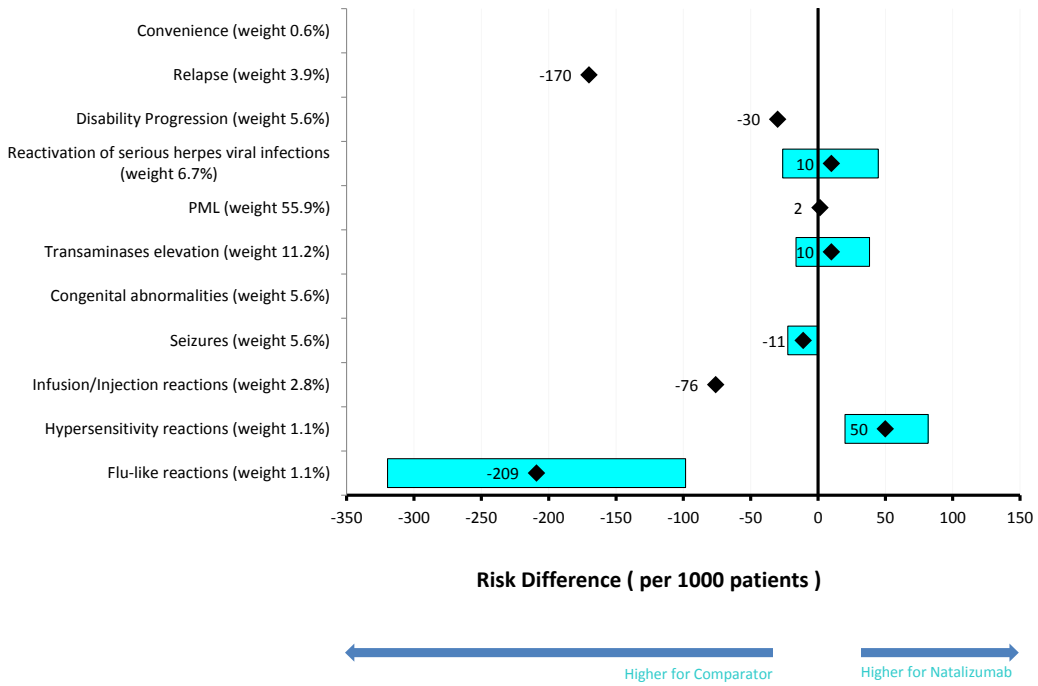


Figure 8: Natalizumab versus interferon beta-1a (Comparator) at time of CHMP re-evaluation



For discussion see section 6.3.2.

5.1.8 Conduction of benefit-risk analysis based on the NNT/NNH concept

5.1.8.1 Methods

The BRAT framework is by intention an open framework that leaves the choice regarding the appropriate quantitative method(s) to be applied for benefit-risk assessment to the user. In this example, the application of the NNT/NNH concept has been chosen that is frequently used in the literature. The definitions given below are taken from Holden et al. (10)..

5.1.8.1.1 Definition of Number Needed to Treat and Number Needed to Harm

The Number Needed to Treat (NNT) is defined by

$$NNT := \frac{1}{(p_1 - p_2)},$$

where p_1 and p_2 denote the proportion of the disease of interest in the control group and the treatment group, respectively.

The Number Needed to Harm (NNH) is defined by

$$NNH := \frac{1}{(q_2 - q_1)},$$

where q_1 and q_2 denote the proportion of the AE (adverse event) of interest in the control group and the treatment group, respectively.

The NNT can be interpreted as the number of patients to be treated in order to avoid one event (e.g. occurrence of a disease) compared to control treatment, while the number needed to harm is the number of patients to be treated in order to cause one additional event compared to control treatment. Assuming that benefit and risk are of comparable medical importance, benefit outweighs risk if NNT is less than NNH.

5.1.8.1.2 Weighted NNT-NNH Comparison

As the considered benefit (e.g. disease avoided) and risk (e.g. adverse event caused) might have different levels of severity, Holden et al. (10;11) suggested a utility weighted comparison defining the weighted NNH_w as:

$$NNH_w := \frac{1}{(q_2 - q_1) * RV}, \; with \; RV := \frac{(1 - utility(AE))}{(1 - utility(disease))}, \qquad (1)$$

where "*utility is defined as the numeric presentation of the patients' preferences for specific outcomes*" with perfect health being represented by a utility value of 1.

With some more detail the Tufts Medical Center defines utility weight as "*the value for a utility or preference for a particular health outcome or health state and can range from zero to one. Utility weights may be measured using direct methods such as time-trade off or standard gamble, or indirect methods such as SF-36, Euro QoL, Health Utility Index, etc.*" (https://research.tufts-nemc.org/cear4/SearchingtheCEARegistry/Definitions.aspx).

It should be noted that due to the applied weighing the weighted NNH_w can no longer be directly interpreted as the number of patients that needs to be treated in order to cause one additional adverse event.

Considering the weighted NNH_w, benefit outweighs the risk, if NNT is less than NNH_w. This is equivalent to

$$\frac{1}{(p_1 - p_2)*(1 - utility(disease))} < \frac{1}{(q_2 - q_1)*(1 - utility(AE))}$$

and to

$$(p_1 - p_2)*(1 - utility(disease)) > (q_2 - q_1)*(1 - utility(AE)). \quad (2)$$

5.1.8.1.3 Generalization of the NNT–NNH Concept to multiple risks and benefits

According to the proposal of Holden (10), multiple risks can be considered in the weighted NNH by summing up the weighted risk differences:

$$NNH_w := \frac{1}{\sum_{i=1}^{k}(q_{2,i} - q_{1,i})*RV_i}, \text{ with } RV_i \text{ as defined in (1) for each } AE(i) \text{ for } i = 1,...,k.$$

In analogy to inequality (2) benefit outweighs risk if

$$(p_1 - p_2)*(1 - utility(disease)) > \sum_{i=1}^{k}(q_{2,i} - q_{1,i})*(1 - utility(AE(i))) \quad (3)$$

The approach already described for the handling of multiple risks can be applied for multiple benefits, too, so that the left hand side of equation (3) can be generalized to

$$\sum_{j=1}^{m}(p_{1,j} - p_{2,j})*(1 - utility(E(j))) > \sum_{i=1}^{k}(q_{2,i} - q_{1,i})*(1 - utility(AE(i))), \quad (4)$$

where the E(1),…, E(m) denote the disease or any other adverse events **avoided** by the treatment when compared to the comparator.

With the definitions

- $p_{j,m+i} := q_{j,i}$ for $j \in \{1, 2\}$ and $i \in \{1,..., k\}$ and
- $E(m+i) := AE(i)$ for $i \in \{1,..., k\}$,

inequality (4) can further be simplified to

$$\sum_{i=1}^{m+k}((p_{1,i} - p_{2,i})*(1 - utility(E(i)))) > 0. \quad (5)$$

The formula one the left hand side can be interpreted as the utility-weighted **net clinical benefit** (NCB) where benefits outweigh risks if the sum is positive.

For the non-weighted case, a definition of NCB is provided, for example, by Sutton (12), where NCB is defined as the sum of expected benefits from treatment minus the sum of expected harms from treatment. When ignoring the weights, this definition is equivalent to the left hand side of (5), because the terms $(p_{1,i} - p_{2,i})$, $i \in \{1,..., m+k\}$, are positive for events related to benefit and negative for events related to harm. The expected benefits as well as expected harms are quantified by differences in proportions that can be interpreted as the percentage of patients with either additional benefit or additional harm compared to control.

Benefit-Risk assessment based on NCB is not limited to the use of differences in proportion, but would also allow other types of data representing benefit and harm like, for example, differences between event incidence rates. However, it doesn't seem to be sensible to mix different types of data for NCB calculation.

5.1.8.2 Results

As shown in the previous section, benefit-risk evaluation following a generalized NNT/NNH concept is equivalent to a **weighted NCB** evaluation. Therefore, NCB is referred to in the remainder of this section rather than NNT/NNH.

For the benefit-risk assessment of natalizumab, two different types of data have been made available

- annualized incidence rates (# events / patient year)
- incidence proportions (# patients with event / # patients).

In order not to mix different types of data in the weighted NCB calculation, the incidence rate for relapses was replaced by the proportion of patients with at least one relapse. This allowed calculation of the weighted NCB according to formula (5) given in Section 5.1.8.1 taking into account the benefits and risk for those quantitative data were available. As a consequence, congenital abnormality was not considered here as no quantification was possible. In addition, convenience was removed from the NCB calculation as this was assessed to be a benefit of only secondary importance. For the benefit-risk evaluation, the so-called preference weights (see Section 5.2.6) were used.

Table 14 Data used for NCB calculation of natalizumab

Event	**Weight**	**Differences in Proportions: natalizumab vs.**		
		Interferon beta-1a	**Glatiramer acetate**	**Placebo**
Disability progression	5.6	0.03	0.07	0.12
Relapses	3.9	0.17	0.21	0.26
PML (at time of approval)	55.9	-0.001	-0.001	-0.001
PML (post-approval)	55.9	-0.0015	-0.0015	-0.0015
Reactivation of herpes	6.7	-0.01	-0.01	-0.01
Transaminases elevated	11.2	-0.01	-0.01	-0.01
Seizure	5.6	0.01	0.00	0.00
Hypersensitivity Reaction	1.1	-0.05	-0.03	-0.05

Event	Weight	Differences in Proportions: natalizumab vs.		
		Interferon beta-1a	Glatiramer acetate	Placebo
Infusion/injection site reactions	2.8	0.076	0.033	-0.056
Flu-like reactions	1.1	0.21	0.00	0.00

Note: Positive differences indicate a benefit of natalizumab vs. control treatment, negative differences indicate a risk.

The numbers given in Table 14 show that, despite the high preference weight, PML will only have little impact on the NCB calculation due to its low frequency. Taking weights into account, PML incidence would need to be 0.012 (or above) in order to cancel out the benefit resulting from disability progression when compared to placebo. To cancel out the benefit in reduction of relapses compared to placebo, a PML frequency of 0.018 (or above) would be required.

The weighted NCB was calculated for the time of submission as well as for the post-approval period. **In all comparisons, a positive NCB was achieved.** As to be expected, the weighted NCB was most pronounced when natalizumab was compared to placebo.

Table 15 Weighted Net Clinical Benefit

	Submission	After Registration
Natalizumab vs. interferon beta-1a	1.041	1.013
Natalizumab vs. glatiramer acetate	1.036	1.008
Natalizumab vs. placebo	1.239	1.211

Notes:

- In this example only the estimate for NCB was calculated. For a proper benefit risk assessment information regarding the variability of the NCB is essential and could be derived by the use of simulation techniques or approximation methods.
- The comparison of natalizumab vs. placebo can be assumed to be the most reliable comparison in this setting, because no head-to-head comparisons of natalizumab with the other active compounds are available. As a consequence, comparative data had to be generated by indirect comparisons requiring additional assumptions (see Sections 4.3.1 and Part 9.1) and resulting in an increased variability of the NCB estimation.

For further discussion see section 6.3.2.

Part 5.2 PrOACT-URL framework using MCDA

5.2.1 Introduction

The PrOACT-URL is a generic framework for framing and analysing decisions described in Hammond et. al (13). This framework has eight steps and we apply it to the benefit-risk context. For the trade-offs step MCDA is used.

Step	Description

1) Problem	Identify the fundamental problem
2) Objectives	Identify the overall value and the criterion categories
3) Alternatives	Identify the possible decisions to be evaluated against the criteria
4) Consequences	What are the observations relevant to the criteria?
5) Trade offs	Trade-off the observations
6) Uncertainty	Deterministic or stochastic sensitivity analysis on the weights and outcome measures
7) Risk tolerance	Are there factors that could affect the decision maker's attitude and accept more uncertainty?
8) Linked decisions	Consistency with other decisions. How this decision could set a president for future decisions.

5.2.2 Step 1: Problem

See section Part 4.1 for details of the problem and section Part 6.1for a discussion of the process

5.2.3 Step 2: Objectives

See section Part 4.1 or details of the objectives and section 6.1.1 for a discussion of the process.

5.2.4 Step 3: Alternatives

Within the context of Benefit-risk analysis the number of possible decision alternatives is expected to be quite small. In this case is it to decide

1. Should natalizumab be given marketing approval at the time of first registration?
2. Should natalizumab be kept on the market given that increased episodes of PML were observed?

5.2.5 Step 4: Consequences

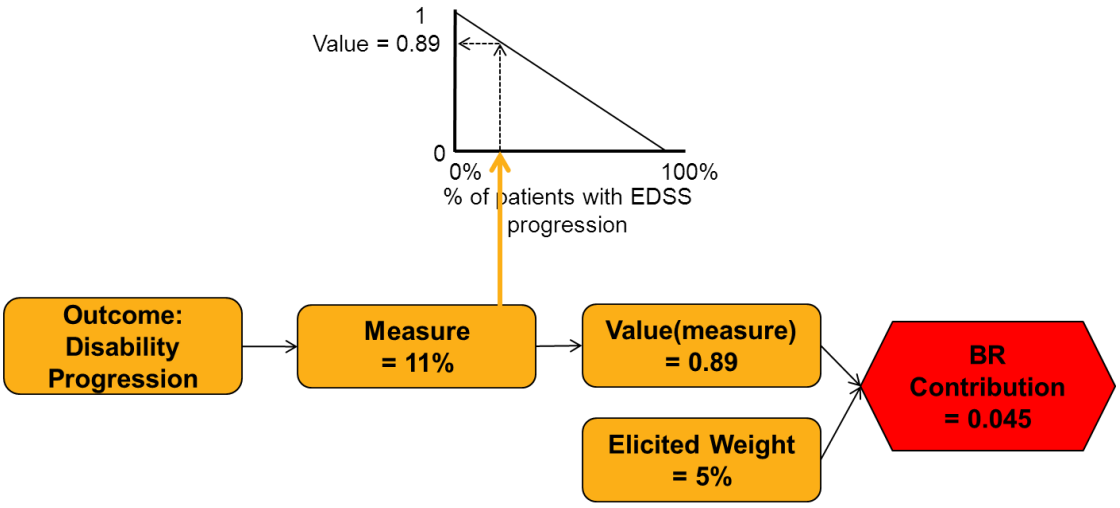
See section Part 4.3 for details of the consequences and section 6.1.1 for a discussion of the process.

5.2.6 Step 5: Trade offs

See section Part 4.4 for details of trade-offs and section 6.1.1 for a discussion of the process. The remainder of this section details the analysis of the benefit-risk analysis.

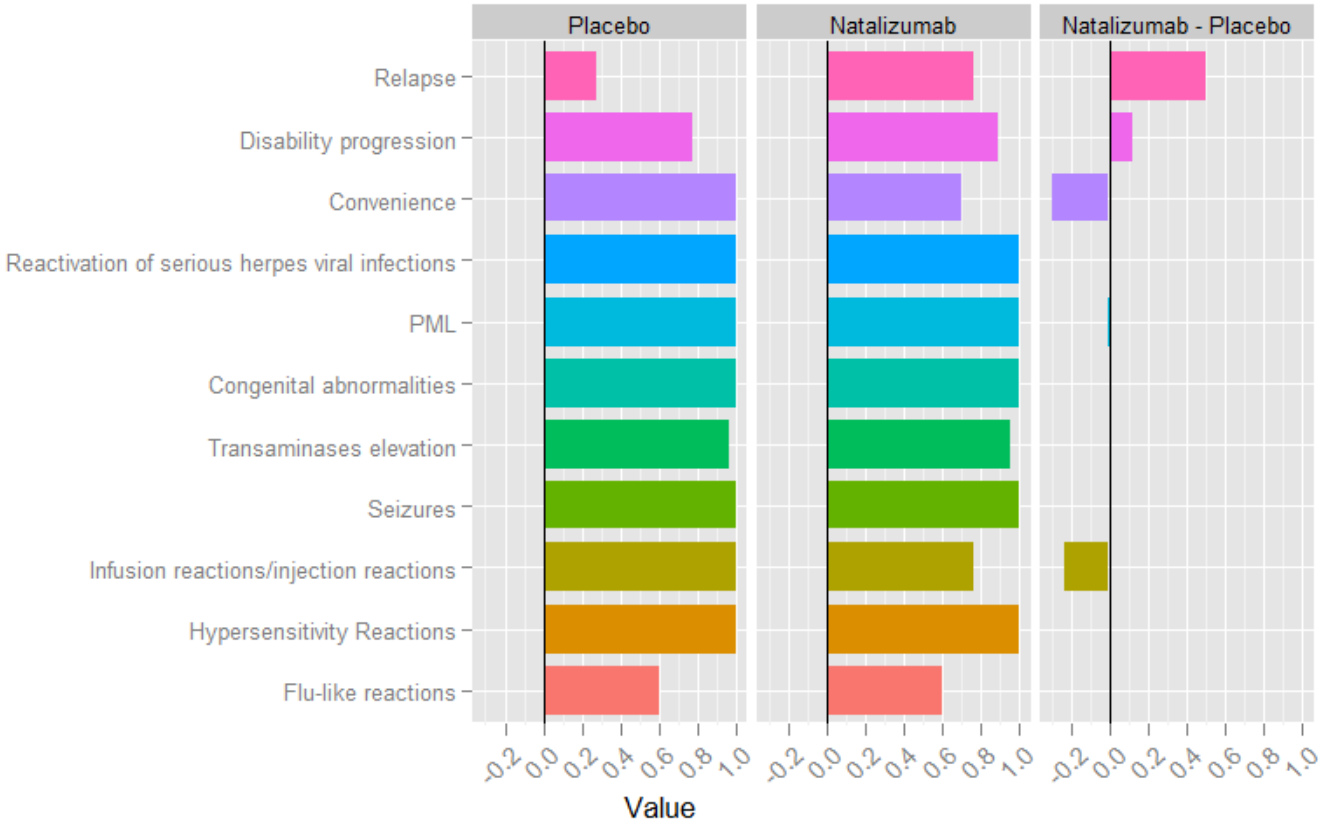
The algorithm for estimating the benefit-risk of a given treatment is straightforward. For each outcome, the observed measure for each treatment has been found in Step 4. This is mapped to a value via its value function and the resulting value multiplied by its weight. For each treatment, this yields the benefit-risk contribution for each outcome. The benefit-risk for each treatment is the sum of these benefit-risk contributions. This is shown in Figure 9.

Figure 9 Benefit-risk calculations



As many of the risks do not occur this leads to a preference value that is 1 for many of the outcomes. This is shown in Figure 10. The issue here is that the places where there is a difference between natalizumab and placebos are lost among the places where there is no value difference. The incremental value between natalizumab and placebo is also calculated. Where there is no difference between the two options, this is zero. It also makes it clear where natalizumab is better than placebo and where it is worse.

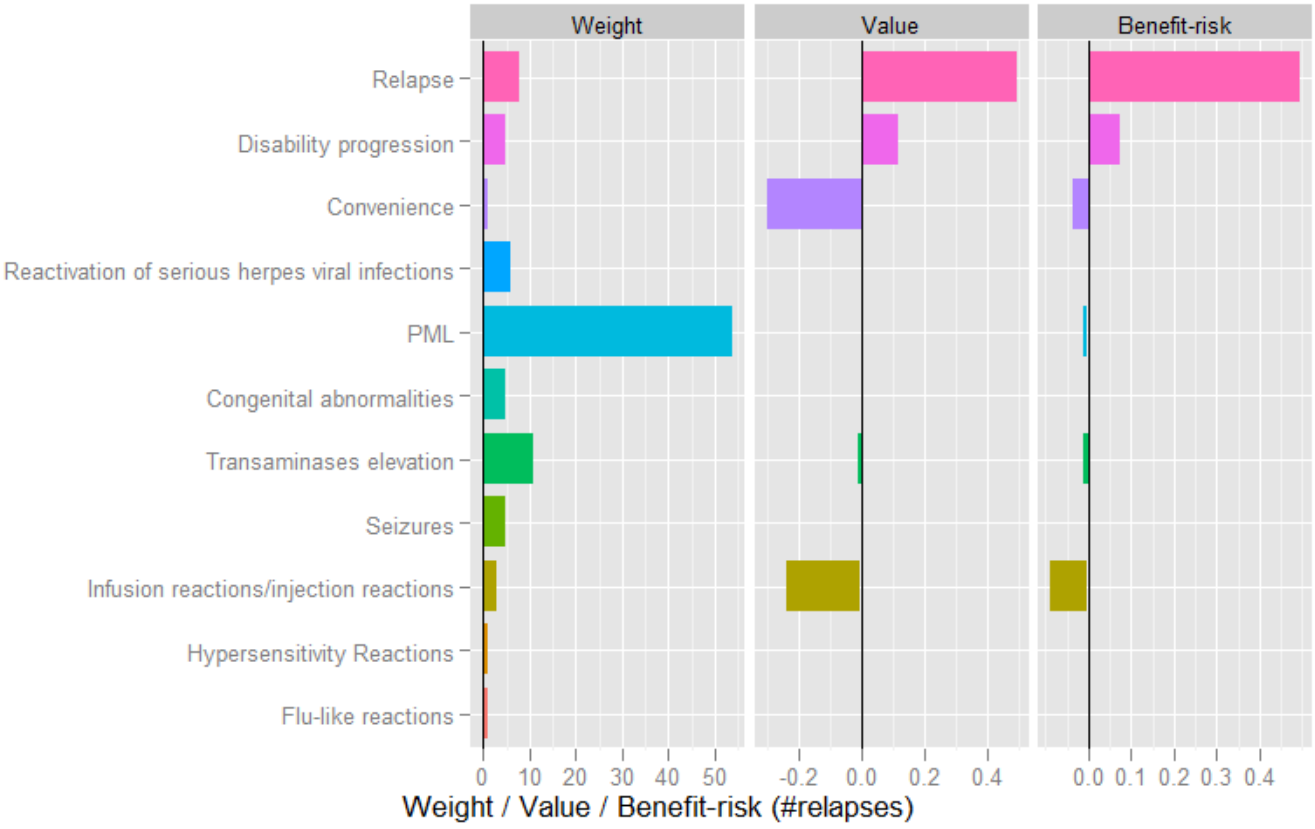
Figure 10 Values for natalizumab and placebo



Next we multiply each incremental value by its weight to find the contribution to the benefit –risk of each outcome. This is shown in Figure 11. As the benefit-risk scale is somewhat arbitrary, we express the benefit-risk in terms of numbers of relapses. This is possible as the value function for the relapse rate is linear. Note that as the value

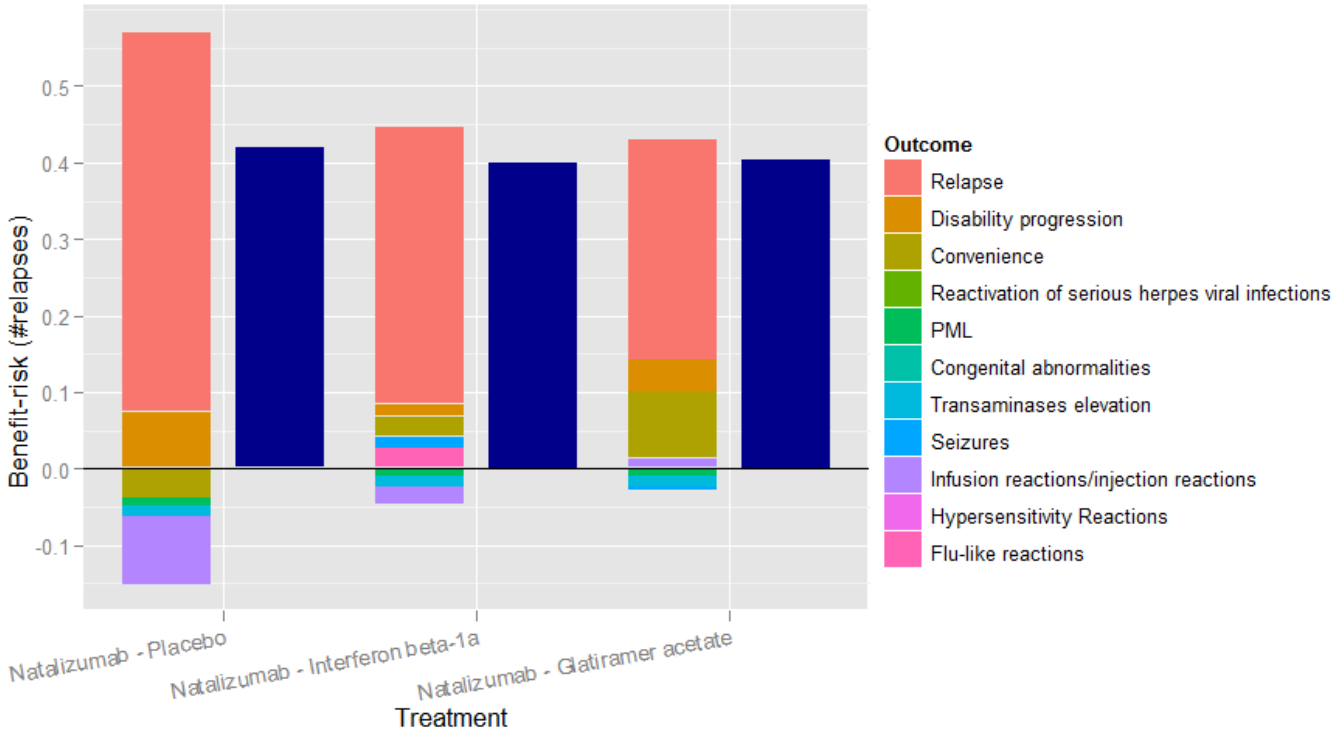
function for the relapse rate has a range of outcome measure from 0 to 2 then a value of 0.5 (say) relates to 1 relapse. From this we see that most of the benefit of natalizumab over placebo is due to a reduction in the relapse rate.

Figure 11 Benefit-risk of natalizumab compared to placebo



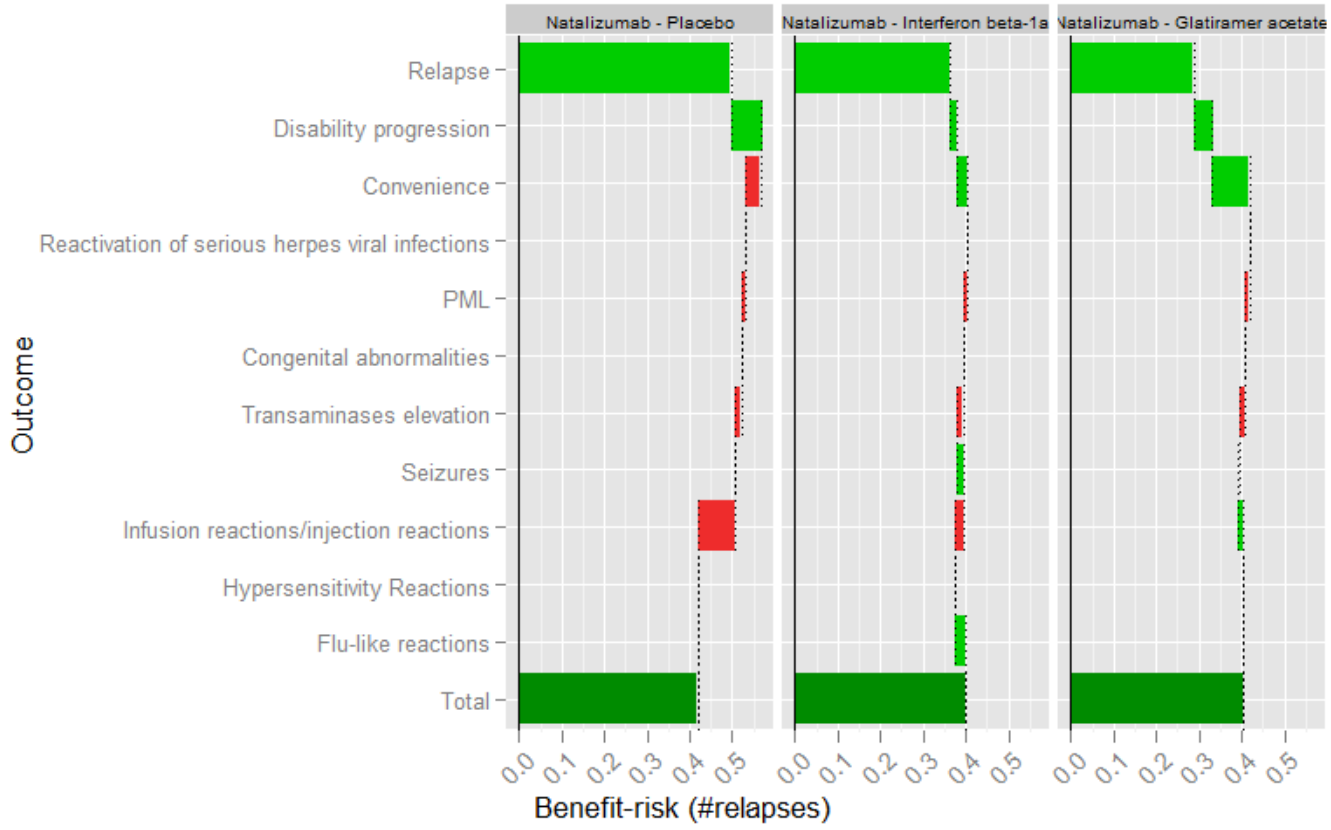
Another way of representing the final panel is as a stacked bar plot. We perform the same incremental-benefit risk calculations for interferon beta-1a and glatiramer acetate compared to natalizumab and stack the positive incremental benefit-risk components vertically above the y-axis, and the negative incremental benefit-risk components below the y-axis. The sum of all the components is also shown as a blue bar in plot.

Figure 12 Incremental benefit-risk of natalizumab compared to placebo, interferon beta-1a and glatiramer acetate represented as a stacked bar chart.



This same information can also be viewed in a waterfall plot. This is like the stacked bar chart in that each incremental benefit-risk component is plotted, with the next component plotted starting where the last one finished. The differences are that the bars are plotted horizontally and are separated from each other. The end of the final bar gives the overall benefit-risk score. The bars are also coloured to show if that component is increasing (green) or decreasing (red) the incremental benefit-risk score. This allows a more direct comparison of components of the benefit-risk between the different treatments.

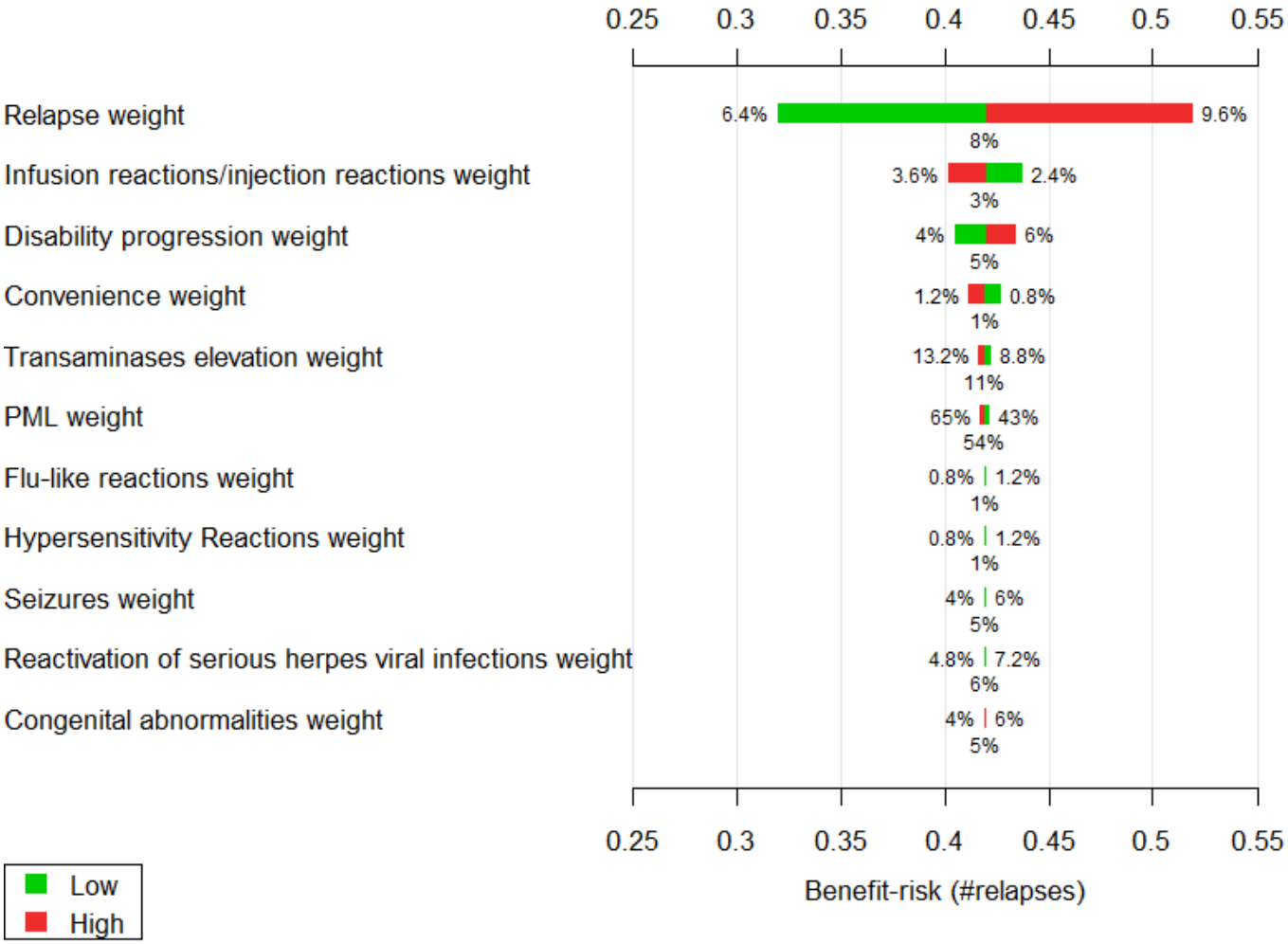
Figure 13 Incremental benefit-risk of natalizumab compared to placebo, interferon beta-1a and glatiramer acetate represented as a waterfall plot.



5.2.7 Step 6: Uncertainty

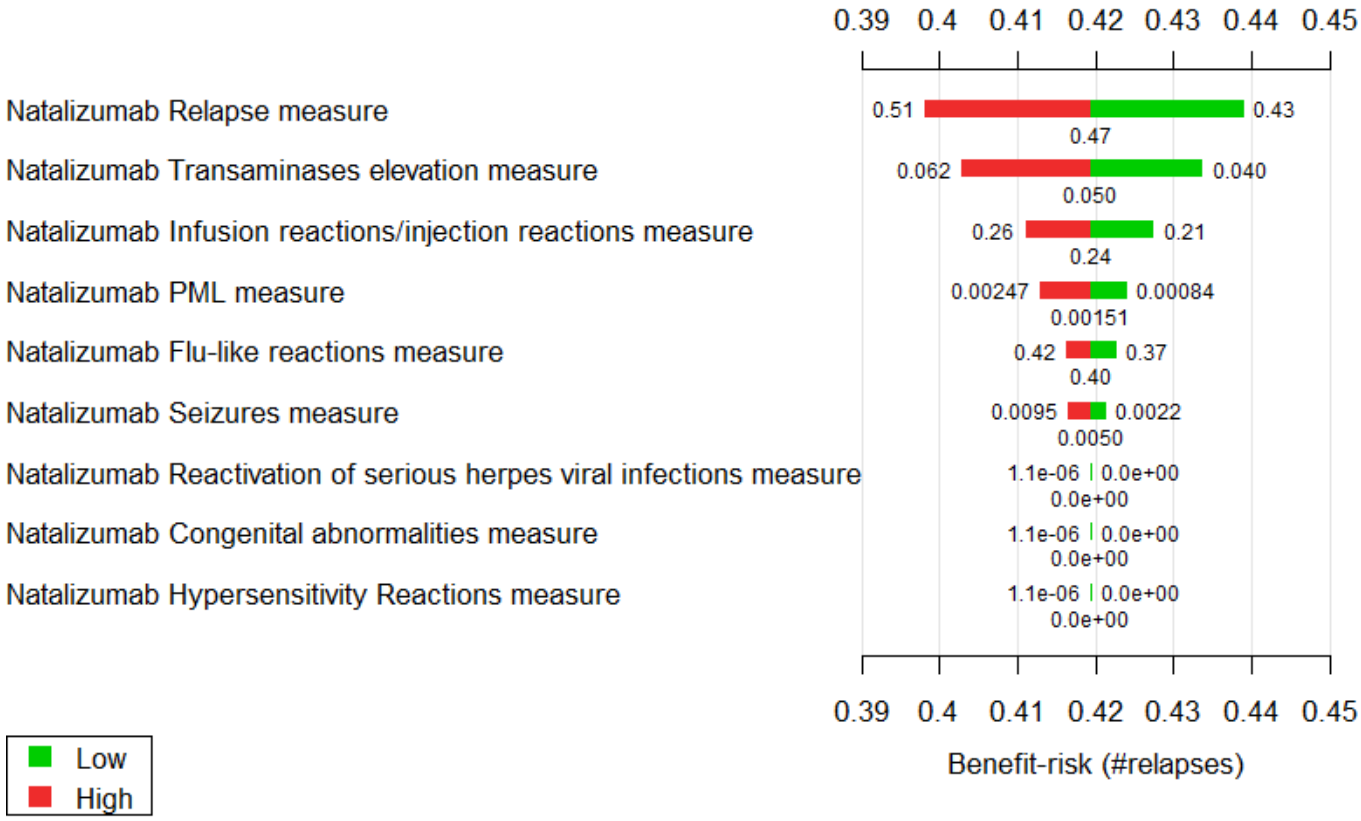
We perform a one way deterministic sensitivity analysis on the weights for all outcomes. This involves changing each weight one at a time by adding on 20% and subtracting 20% off the relative value, then assessing how this changes the incremental benefit-risk of natalizumab compared to placebo. Weight is a subjective measure and 20% is a nominal choice for the range for the sensitivity analysis. This is shown in a Tornado diagram in Figure 14. The base case value of the weight for each outcome is shown under the middle of each bar. The low and high values of each weight are shown at the ends of the bars. Low values are shown at the green end and high values at the red end. The incremental benefit-risk at the base case is the x-axis value at the middle of all the bars, and how this changes with each weight is shown by the position of the bar ends. From this plot we see that changes in the weight of relapse has the most influence on the benefit-risk score.

Figure 14 Tornado plot to assess how changes in the weights affect the incremental benefit-ris score.



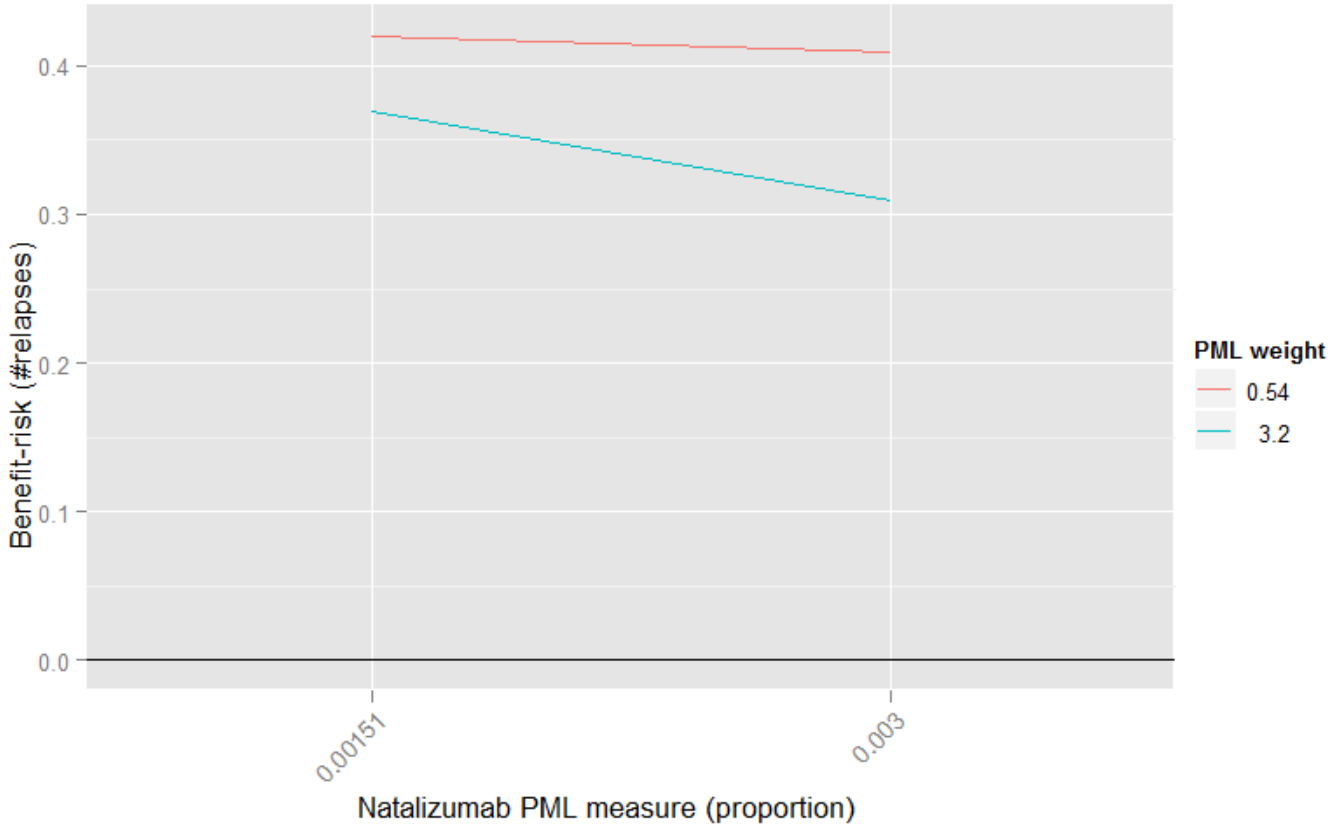
A similar one-way deterministic sensitivity analysis is performed for the measures for each outcome for natalizumab. The measures are varied one at a time using the 10% and 90% quantiles of the distribution of the estimated value. This distribution is not known for the disability progression so it is fixed at its mean value. This figure shows that the outcome having the most effect on the incremental benefit-risk is the rate of relapse when given natalizumab.

Figure 15 Tornado plot to assess how changes in the natalizumab outcome measure affect the incremental benefit-risk score.



As PML is a key issue, a two way sensitivity analysis is performed on this outcome. Both the weight given to PML and the proportion of patients developing PML are varied at the same time. This is shown in Figure 16. The x-axis shows the proportion of patients developing PML. This starts at the base case values of 0.151% and is increased to twice this value to 0.3%. If one looks at the EDSS progressions prevented by natalizumab, and the PLM rate at the time that natalizumab was withdrawn from the market, the implied weight given to PML compared to EDSS progression is 3.2. Each line in the Figure denotes the weight given to PML, this starts from the elicited weight of 0.54 and is increased to this implied weight of 3.2. In the sensitivity analysis the weights and incidence are only increased to "stress test" the base case analysis and find the situations where the benefit-risk is no-longer positive. In the elicitation process the weights are normalized to sum to one, this is not necessary for this sensitivity analysis. In effect the weight given to PML is changed from being equivalent to 10 EDSS progressions to being equivalent to 60, and all other outcomes keep their same relative weights. The y-axis shows the incremental benefit-risk compared to placebo. This figure shows that at the base case natalizumab has a positive benefit-risk balance, and this is robust to both doubling the rate of PML and increasing the weight of PML six-fold (to the implied weight that lead to the withdrawal of natalizumab).

Figure 16 Two way sensitivity analysis plot, showing how changes in both the number of patients developing PML and the weight associated with PML affect the benefit-risk score.



5.2.8 Step 7: Risk tolerance, Step 8: Linked decisions

These two steps are not considered relevant for this case study as we consider this a standalone example to explore the methodology.

Section 6 Discussion

Part 6.1 Methodology: Appropriate frame

Both methodologies were based on the same frame. This is described in section Part 4.1

Table 16 Assessment of appropriate frame for benefit-risk approaches through practical experience

	Comments	Proposed improvements and/or extensions
Rational for the chosen approach	The BRAT and PrOACT-URL essential given the same guidelines for defining the decision context. We found this fit for purpose. The group comprised experts from clinical, drug safety, epidemiology, statistics, and Health Economics. This was facilitated by a member of the team.	
	The hardest issue to deal with was the time horizon for outcomes. A two-year horizon was chosen as this is the duration of the pivotal clinical trials. However, this may not capture rare events or events that take longer to manifest. It was assumed that occurrence peaks at 24 m of treatment (Kappos (5)). PML is not observed in Monotherapy natalizumab, so the data sources were extended to post-marketing data. The time frame could not be precisely found from available data, but it was reasonable to assume that it was about two-years. Regulators decision was based on the assumption that there is not much more efficacy to achieve with the available drugs and a high medical need remains as outlined in CHMP Scientific Discussion (see above)	How to deal with outcomes on different time horizons? Should discounting be used?

	Availability of additional risk minimization measures such as the JVC antibody test recently developed by Biogen would make a difference in benefit risk assessment but will not be considered in this exercise due to the paucity of information currently available in the public domain, and also was not available at time of the CHMP reevaluation in 2010.	
	Standard information in SPC on frequencies was considered more vague (PML uncommon, between 1 in 100 and 1 in 1000)	
	Perspective taken was the regulator, assuming they are a perfect agent for the patient.	Reasonable to include as many perspectives as possible at this stage so you do not have to reopen the discussion later (ie.e. agree on this stage whether the relevance for Physicians **and** Patients should be taken into account). However, it is important to always be clear which perspective is being taken, and multiple perspectives would make the whole process more complicated.
	Population was patient with RRMS.	While it was recommended by the BRAT user guide to look into further differences between the patient populations treated in the studies when initiating comparisons across trials, the Team did not look into further potential differences of comparators such as CIs, dosing etc as the first level (indication) already needed so many assumptions. When taking real world decisions, attention should be given to all key differences between the patient populations and those should be carefully documented.

Implementation	Initially Xmind, a mind mapping tool, was used to facilitate this step. This step was performed by the PrOACT-URL team in a meeting lasting two hours. BRAT framework guides well through the process, the decision context table is easy to use. Considers many important aspects comprehensively. In fact, the BRAT guidelines were used to influence this step in the PrOACT-URL framework.	
	Two time points were chosen. The decision at the time of registration and post-market decision point.	The analysis performed could also be used for post marketing risk management. When events are observed they can be include. However, it is important to know what the time frame and the person-years at risk are for such an analysis.
Making comparisons	Both frameworks allow for comparisons. What is the correct comparator for decision making? Differences in indications / trial patient population: Interferon beta-1a/glatiramer acetate is approved in less severe patients (first line) than natalizumab (second line). The CHMP Scientific Discussion (see above) revealed significant benefits for natalizumab in a subgroup analysis in a population with more advanced MS . This analysis obviously served as a basis for the EU approved indication. Ideally we would have data on the comparators interferon beta-1a and glatiramer acetate for second line therapy but unfortunately the data are not reported for this subgroup in the available sources. Patients in the Phase III studies were mostly treatment naïve. It should be assumed that natalizumab will be used as indicated only when other drugs have been used before. So the natalizumab Benefit risk balance would have to be better than the benefit risk balance of interferon beta-1a or Copaxon rather than comparing it to placebo, taking however into account that interferon beta-1a and Copaxon may have less remaining efficacy at this stage of the disease.	

	As long as those data on remaining activity of interferon beta-1a and Copaxon in these more advanced MS patients are not available one could also assume in this model that the benefit of both is going down from original Phase 3 results and placebo may be therefore be the better comparison. However, this cannot be verified based on available data and the assumption may be wrong. Patient perspective: patients would prefer drug not placebo, would prefer natalizumab assuming it works as approved and if their old drug does not work any more"	
Limitations	The framing is a very open guideline which did not put limitations on what could be done.	

6.1.1 Meaningful reliable information

Both methodologies were based on the same objective and subjective data. This is described in section Part 4.3 and Part 4.4.

Table 17 Assessment of using meaningful reliable information for benefit-risk approaches through practical experience

	Comments	Proposed improvements and/or extensions
Rational for including outcomes	The risk management plan was the starting point for the risks and the common primary and secondary endpoints the starting point for the benefits. Outcomes were excluded if • They are surrogates for other endpoints in the value tree. • They were very similar to other included outcomes. • They were thought not to be relevant form the regulatory perspective • They were easily manageable. PML had been a topic of discussion at time of approval in EU, so the PML risk was also prominent in the Benefit risk tree at time of approval	
	Consideration of other benefits There was a general concern that the benefit risk tree will provide visual imbalance that may transfer into numerical differences later if the group is too selective on either benefits or risks. If an outcome is missing from the tree then it is the same as giving that outcome no weight. Also there is a tendency to focus on risks instead of benefit. On the other hand, the risk of double counting must be avoided. We believe all benefit outcomes relevant for the patients should be considered and this would also include amongst others. Reduction of T1 and T2 lesions, life expectation, quality of life, black holes. Unfortunately the pivotal trials did not deliver comparable data on all these for all treatments, so data availability was a driver of the result.	
	BRAT framework: As the convenience of drug administration is not primarily a medical benefit, it was ignored in the quantitative benefit risk evaluation in order to avoid overweighting. The latter might occur in the quantitative evaluation as the convenience advantage of a compound affects all patients. In this particular case convenience is not a key differentiator between the compounds.	

	In this specific case it was agreed to keep the value trees the same when comparing to Copaxon and interferon beta-1a. It was discussed however that the distinction between benefit and risk in the value tree is in general a qualitative definition that may change when using quantitative comparison with alternative treatments. I.e. depending on the active comparator, some risks of natalizumab might become benefits in the Net Clinical Benefit (NBC) evaluation. While there is no need to modify the decision tree, such changes in perspectives should be documented when calculating NBC.	
Scale of measurement	For benefit risk evaluation, absolute risk differences provide more transparent information than relative differences like odds ratios or relative risks. While the latter can provide a good first impression comparing treatments regarding one benefit (or risk), absolute differences are more appropriate for the comparison of more than one item. Absolute numbers are also needed for the value elicitation.	
Data	For the benefits, data were available as they were primary of secondary endpoints. However, uncertainty may not have been reported (only p-values). For risks, they were sometimes available in the literature. Where they were missing, the registration documents from the FDA were used. We assumed that absence of evidence meant that no events occurred. However, there could be under reporting bias here. Needed to assess if difference in definitions used were comparable, and that the same definitions for risks were used. The frequency of PML is difficult to determine from the varied sources. EMA publications and a recent review by Kappos et al, 2011 were selected as primary source. Zero PML incidences were assumed for interferon beta-1a, glatiramer acetate and placebo at both evaluation time points. Variability of source data and subsequent transformation: Different experts will select different data.	Original data plus modified data (e.g. allowing indirect comparison across trials) must be well documented and provided together with the result.
Clinical	Time consuming to find consensus. 2 x 2 hour meetings.	Start with a proposal already prepared, then get consensus in a group – a consensus meeting of all functions is a

judgements		prerequisite for finalization of the value tree.
BRAT framework for value tree	This was used for the process. Easy to use, well supports the creation and flexible adaption of the benefit risk tree. The terminology of the BRAT guidance documents is more focused on industry terminology not designed at this step as a regulators tool. Regulators would expect the data sources being made available by industry and as part of the dossier. Would need some minor modifications to the guidance for a more neutral use	
BRAT tool for data	BRAT provides a summary table in an electronic excel tool that allows to enter all data in a predefined format as source table for further graphs and overview tables (for output see below). The tool is available free of charge from the BRAT Team. It is easy to populate and uses any data available Deals well with missing data and highlights them in a control step Tool allows entry of data in all different formats, i.e. in percent or annualised rates. Provides good flexibility on what is presented, allows to enter all data in all forms you want to use (Odds ratios, absolute differences etc.). At the same time the data source table in the excel tool itself provides limited guidance which data to select, so it should not be used in isolation. The BRAT framework and in particular the BRAT users guide should always be consulted.	
Decision making	The mild increase of occurrence of a very rare serious event in the context of several well documented benefits and more common risks makes it difficult to balance from the data alone. Under these circumstances, weighing the risks seems a prerequisite.	

6.1.2 Clear values and trade-offs

Table 18 Assessment of the availability of clear values and trade-offs for benefit-risk approaches through practical experience

	Comments	Proposed improvements and/or extensions
Judgement values	BRAT allows, but does not insist on preference values. PrOACT-URL is more explicitly require the use of preference values. Judgements were obtained from patient representatives using bottom-up swing weights. Given the nature of the problem – balancing one serious and rare adverse event over several clinically relevant and frequent benefits – the team agreed that weighting from patient perspective is necessary as basis for decision making. Usually Regulators decide on behalf of the patient while risk tolerance of Regulators may be much lower than that of a patient. So a carefully elicited patient input is required. Such preference elicitation / weighting needs a thorough methodology Questionnaires may provide very different values depending on patient experienced with one of the events. I.e. EDSS does usually not follow a linear value function as avoiding a wheelchair has a high value for the patients. Even if science does not provide a basis for this, risks may be perceived to occur "more to others" and later in time. This may lead to trading off early benefit versus late risks. "I would like to play with my child today and risk PML in a couple of years". While this was only transferred from hearsay, it is clear that those patients who are asked for their	As some of the weights provided were questioned the team recommended a sensitivity analysis, in particular related to the two main benefits versus PML risks, in particular looking at higher avoidance preference for the life threatening risk of PML. Scale bias: there are cognitive biases in realistically understanding large numbers. Patients tend to put more value on avoiding harm that they are experiencing, and may underestimate the harm that they could potentially experience. Not having experienced a harmful event may lead to a further underestimate of the preference for avoiding it The tools and research in this area needs to be further analysed in the context of Benefit Risk assessments by Regulators.

	weights need to well understand the risks and benefits they weigh on before they do. Unguided questionnaires provide the same information to each patient and thereby might avoid bias, while interviews may be preferable to fully explain and discuss the disease and the risks involved. To overcome this dilemma. So called “Decision Aids” have been proposed as tools to improve patient´s decision making in presence of a Physician [Col, Shared decision making , Communicating Benefits and Risks: An evidence based users guide, FDA August 2011, Ch 17, pp173-183: This step provides an opportunity for stakeholders with different perspectives to discuss their judgment about the priority of the outcomes and the rationale behind those judgments. The clear visibility of differences among stakeholders and the reasons for them can foster understanding of varying points of view and elucidate how various stakeholders may draw diverse conclusions. [The PhRMA BRAT Framework for Benefit-Risk Assessment User’s Guide to the Process: Final February 4, 2011 The Benefit-Risk Action Team] The elicitation of reasonable preference weights most likely depends on the stakeholder (e.g. regulators, physicians, patients) involved. Therefore sensitivity analyses using preference weights from different groups of stakeholders are recommended to check the robustness of the decision. For BRAT very general guidance given. No preferred tool recommended. Also PROACT URL does not specify a tool.	
Definition of benefits and	It is important to have consistent definition between the value tree, the objective data	

risks	extracted and subjective data elicited. A broad definition makes data extraction easer, but preference weighting harder. A narrow definition makes data extraction harder, but preference weighting easier.	
Common scale	PrOACT-URL using MCDA outcomes are mapped to a value scale. BRAT Uses a risk difference scale. So outcomes must be in terms of events.	
Results interpretable	Adaption of forest plots to allow display of risks and benefits of very different occurrence was not immediately successful: Both visual displays were not sensitive to cover uncommon or rare events in the presence of common or very common events. Even if common minor ADRs (flu like symptoms) were excluded as they are dilutive to the overall picture, the overweight of frequently occurring benefits dominated the visualization. The team referred back to the original questions and compared the displays. Trading off early benefit versus late risks was discussed; theoretically the risk of later occurrence of a PML could be discounted versus an earlier benefit from the drug. The team discussed this and discarded the argument - the risk of PML seems highest at/after 24 m (Kappos et al 2011) while the study results showed increasing benefit over 24 m (see CHMP Scientific Discussion). Consequently from a scientific point of view, time should not be major driver of preference.	This visualization problem is a) not specific to the forest plot and b) might be mitigated by re-scaling (e.g. log transformation). However, when using a transformed scale the resulting figures have a higher risk for misinterpretation as not everyone involve in BR might be familiar with log scales. Include in this discussion a qualified statistician who knows the data and a qualified clinician who understands the disease area and the clinical implications of the findings What is the expected impact on the patient? Reconsider all uncertainties and assumptions as outlined above: Completeness and quality of the studies, the quality of the data, the potential for bias and confounding in the included studies, appropriateness of analytical methods in the included studies and in the Framework Process,

	The team is accepting the assumption that all treatments are equally effective at time of second line treatment PML is so rare that it does not obviously outweigh the clinically relevant benefit with natalizumab PML as a rare event does neither show up predominantly in the table nor the forest plot Even if data are objectively displayed, the subjectivity of decision making does not completely go away as several criteria regarding benefit and risk have to be weighed against each other, so that appropriate preference elicitation is the most critical parameter in the process. The framework makes this subjectivity explicit and thereby increases transparency. The BRAT user guide provides advice on the evaluation step.	appropriateness and relevance of selected outcomes and measures used in the Framework, the magnitude of the treatment effects, accepted or established clinically important levels or changes in health states represented by the outcomes and measures, and the importance of the outcomes and the treatment effects to health care providers and patients.
BRAT	The Framework provides a structured assessment process that is transparent all relevant information needed for discussion of the benefit-risk profile for a product an aid to decision-making, but not an interpretation or conclusion in isolation A decision preparation tool for decision by people not in depth involved in all data, committee, high level	

	Displays: an easily accessible and readily understood summary data displays and underlying data in an organized fashion. The key Benefit risk table provides good overview on one page on all benefits and risks, even more so if preferences added. Overall, accumulating the data in this specific table is additional work if the Project Teams of the MAH have already worked out other displays. One standardized approach will however allow comparison across portfolios so a standard display does add value. Provides the information more intuitive and for visually oriented people BRAT excel tool All steps in one file – one source for discussion, depending on variety of data types not all wishes can be fulfilled, good for those workshops devoted to the exercise due to quick on line adaption by moderators – place the data in different scenarios Highlighting of risk priority/importance, automated severity ranking or ranking acc to weights would be helpful The forest plot display does not immediately allow weights to be added. Benefits and risks could however be ranked in order of preference with the weights added to the label. Weights should also be added to the benefit risk value tree in the next version.	
Weighted Net Clinical Benefit	Limitations Both the MCDA and the weighted NCB do not directly allow adequate consideration of any changes in risk over time. However, this can be done by focusing on various time periods of interest. This approach can also be applied in order to deal with various	

	subgroups of interest that are assumed to have different benefit risk characteristics. In order not to mix different types of data, only differences of incidence proportions were considered here although event incidence rates might provide more relevant information especially for recurrent event like relapses. Nevertheless, it would have been possible to directly use the relapse rates in the quantitative benefit-risk assessment, but this would have implied that two relapses seen in one patient have the same weight in the assessment as two patients with one relapse each. Depending on the endpoint, this might be a reasonable assumption, but should be checked individually. Another alternative would have been the consistently use of event incidence rates for all beneficial or detrimental events under consideration. However, adverse event data are often not reported this way in the literature, so that individual patient data would have been required in order to calculate these rates.	

6.1.3 Logically correct reasoning

Table 19 Assessment of the logically correct reasoning for benefit-risk approaches through practical experience

	Comments	Proposed improvements and/or extensions
Form of data	BRAT with w-NCB needs outcomes to be expressed as events. PrOACT-URL with MCDA: can accept any form of data. Outcomes should be expressed on an absolute scale as this is the scale that preference values are expressed.	Continuous or ordered categorical data could be dealt with by definition of either one or a set of appropriate dichotomous variables (e.g. responder definition in case of continuous efficacy data).
Uncertainty	BRAT with w-NCB Risk difference are presented with CI. In principal accounting for deterministic and stochastic uncertainty is allowed in the method, but was not implemented. PrOACT-URL with MCDA: Uncertainty is an explicit step in the process. Accounting for deterministic and stochastic uncertainty is possible, and one way deterministic sensitivity analysis on the weights and measures was applied.	Uncertainty of the weighted NCB approach can be assessed by exploring the variability using, for example, approximation techniques. Further sensitivity analysis can be performed to explore showed at which points the decision made would be changed,
Theoretical justification	The data synthesis assumed that there is no covariate-treatment interaction, so that the absolute effect rate can be found from multiplying the relative effect by a common placebo rate. Both frameworks were based on utility theory, assuming mutual preference	

	independence of criteria and coherence of the elicited preference values.	
Technical flaws	Scale bias: there are cognitive biases in realistically understanding large numbers. Potential problems with the hierarchical weighting process with respect to scale bias. Injection site reactions and EDSS are not compared directly. Injection site reactions are compared to PML, and EDSS progression is compared to PML. The common comparator is potential subject to scale bias, and the indirect comparison of injection site reactions and EDSS (1 patient experiencing EDSS progression is worth 2 patients experiencing injection site reactions) may be difference from if the question had asked directly. Patients tend to put more value on avoiding harm that they are experiencing, and may underestimate the harm that they could potentially experience. Not having experienced a harmful event may lead to a further underestimate of the preference for avoiding it	

6.1.4 Commitment to action

Table 20 Assessment of the commitment to action for benefit-risk approaches through practical experience

	Comments	Proposed improvements and/or extensions
Develop insights	Both methods unpack the main drivers of the decision and highlight how these differ between treatment options.	
Relevant to decision	The final results of both methods directly assess the benefit-risk balance, which is the bases on which decisions about registration are made.	
Easily communicable	The BRAT/ w-NCB method is more easy to communicate than the PrOACT/MCDA method, however, both can be readily understood, and both are transparent.	
Audit trail	Both methods leave an audit trail, allowing the methods to be replicated using data in the public domain.	
Recommend approach	Both approached are recommended. If the decision context allows all outcomes to be expressed at events the w-NCB approach used in the BRAT/ w-NCB has the advantage of being simpler. If it is required to use a richer set of outcomes on different scales the PrOACT/MCDA method would be needed.	

Part 6.2 The assessment of benefit-risk balance

6.2.1 BRAT Framework using weighted Net Clinical Benefit.

Decision 1: Should natalizumab be given marketing approval at the time of first registration?

PML is a rare serious event to be balanced versus a clearly documented benefit on relapses and disease progression. The decision is mainly driven by patient preference and benefit risk balance seems positive to the team on the basis of the elicited weights by the patient groups.

Decision 2: Should natalizumab be kept on the market given that episodes of PML are observed at the time that these episodes were observed (at time of CHMP re-evaluation)?

Though the occurrence of PML increased with more and longer exposure, the overall picture remains the same.

6.2.2 PrOACT-URL framework using MCDA.

Decision 1: Should natalizumab be given marketing approval at the time of first registration?

The PrOACT-URL analysis used the incidence of PML observed at the time of CHMP re-evaluation. This lead to a benefit-risk balance in favour of natalizumab compared to all the other treatments. As removing the PLM incidence would improve the benefit-risk balance, we conclude that natalizumab should have been given marketing approval at the time of first registration.

Decision 2: Should natalizumab be kept on the market given that episodes of PML are observed at the time that these episodes were observed (at time of CHMP re-evaluation)?

The incidence of PML at the time of CHMP re-evaluation still leads to a benefit-risk balance in favour of natalizumab compared to all the other treatments. The sensitivity analysis (Figure 16) shows that this balance is robust to changes in the weight given to PML or the incidence of PML.

Part 6.3 Visual representation of benefit-risk assessment results

6.3.1 BRAT

The displays without further condensation delivers benefit-risk information for decision making without drawing final conclusions or synthesizing to one number. Such display is considered to be acceptable and easily interpretable by decision makers without sophisticated statistical expertise.

Initially the display of a benefit risk tree, absolute numbers in tabular forms and relative risks as forest plot are recommended as examples. Other displays and other methods can be easily integrated in the framework at any point in time.

An issue with displays used within the BRAT visuals representations is that rare outcomes shown with common outcomes can be misleading. The rare events may appear not to occur as they are beyond the resolution of the plot. The seriousness of the event is not captured by the plot and this could be misleading.

6.3.2 PrOACT-URL

The PrOACT-URL does not prescribe specific plots, but does emphasize the importance of communication of results and sensitivity analysis, and the plots selected are in this sprit.

As many of the risks did not occur, this lead to a preference value that was 1 for many of the outcomes. This is shown in Figure 10. This lead to the places where there was a difference between natalizumab and placebos being

lost amount the places where there is no value difference. We found it best to display the incremental value between treatments.

Benefit-risk analysis has a lot of moving parts and it is helpful to not just give the overall benefit-risk figure but also open up the machinery and expose how this overall figure is constructed. Benefit-risk, and decision analysis in generally is not about calculating the right answer, but giving clarity to decision makers. Opening up the analysis in this way can help elucidate the decision and also build confidence with the decision maker to the appropriateness of the analysis.

It was helpful to break down the overall benefit-risk into the contributions from the different components from each outcome. This was done in three ways.

- Figure 11. Horizontal bar chart, each bar giving the incremental benefit-risk components, and also the weights and values that make up the comments.
- Figure 12.A vertical stacked bar chart, with positive incremental benefit-risk components above the x-axis and negative ones below, with a total benefit-risk bar plotted next to the stacked bar. With a pair of bars for each comparison.
- Figure 13. A waterfall plot, which is similar to the horizontal bar chart, except that the end of the previous bar determines the start of the next bar, so the end of the last bar gives the overall benefit-risk. This is like a hybrid of the other two plots.

The sensitivity analysis was represented by one-way and two-way deterministic sensitivity analysis plots.

- Figure 14. Tornado plot is used for representing the one-way sensitivity analysis which shows the effect of varying each outcome weight one at a time. This shows which weights have the most influence on the benefit-risk.
- Figure 16. The two-way sensitivity analysis was useful for focusing in on the key issue of PLM, and varying both the weight and incidence together, and assessing in which situations the incremental benefit-risk becomes negative. This highlighted that the incremental-benefit risk is sensitive to changes in the weight and incidence.

There were issues of how to represent numbers on very different scales on the same plot. Very low weights or rare outcomes may not be resolved on a plot.

Section 7 Conclusion

Both the methods used in this case study come to the same conclusion that the benefits of natalizumab outweigh the risks. This is the case both at the time of the time of submission and time of the CHMP re-evaluation. The sensitivity analysis performed using the PrOACT-URL framework demonstrated that this conclusion is robust.

Both methods used the same objective and subjective data sources. The BRAT framework using the weighted NNT/NNH approach, can be shown to be equivalent to a weighted net clinical benefit, is very similar to the PrOACT-URL framework using MCDA, so it is not too surprising that the conclusions of the two methods are similar.

The two frameworks are essentially the same, with an almost one to one mapping between the different steps in the process. The explicit iterative step in the BRAT framework is especially helpful, as iteration is needed in this process and making this explicit is an advantage. These frameworks alone are helpful for structuring and giving insights to the benefit-risk balance. Using them defines value trees, summary tables and plots are useful even if the quantitative aspects of the frameworks are not performed.

The BRAT framework emphasizes the value tree (criteria tree) build-up, data selection, data preparation as well as tabular and graphical data presentation. However, as no tools for further aggregation of data are provided, only a separate assessment of each benefit and risk as well as first visual comparisons are possible. This was consciously proposed to avoid synthesis of data into sophisticated statistical models which may not be easily understandable [The PhRMA BRAT Framework for Benefit-Risk Assessment User's Guide to the Process: Final February 4, 2011 The Benefit-Risk Action Team]. Furthermore, several methods supporting benefit risk assessment are available and the choice of the most appropriate might depend on the specific scenario under consideration. The Excel tool is designed to support pharmaceutical companies in collecting all available and relevant evidences of a new drug for communications with regulatory authorities. The use of such a framework can increase transparency, predictability and consistency with which benefit-risk assessments are conducted. This is an easy-to-implement tool to structure simple decision problems on daily basis.

The use of NNT/NNH has been criticised as it does not allow different outcomes to have different amounts of importance. The development of this to the weighted Net Clinical Benefit overcomes this concern of the NNT method. The weighted Net Clinical Benefit approach is a special case of MCDA assuming that all benefit/risk criteria can be quantified as rates/proportions and linear value functions are used. The Net Clinical Benefit method is arguable easier to explain and to understand than MCDA, which is an advantage, and in many cases this advantage could be worth the decreased flexibility.

A key concern with the frameworks is the elicitation of importance weights for the different clinical outcomes. It is difficult to put a preference weight on a rare serious event. Firstly, could be due to scale bias: there are cognitive biases in realistically understanding large numbers. Secondly, patients tend to put more value on avoiding harm that they are experiencing, and may underestimate the harm that they could potentially experience. Not having experienced a harmful event may lead to a further underestimate of the preference for avoiding it.

Furthermore these issues are compounded with use of the hierarchical weighting process. For example Injection site reactions and EDSS are not compared directly. Injection site reactions are compared to PML, and EDSS progression is compared to PML. The common comparator is potential subject to this scale bias, and the indirect comparison of injection site reactions and EDSS (1 patient experiencing EDSS progression is worth 2 patients experiencing injection site reactions) may be difference from if the question had been asked directly.

PML is “non-influential” as it is too rare to be significant, whereas the nuisance risks of injection reactions, although of low weight, are much more prevalent, and have a bigger influence on the benefit-risk. This combination of rare serious and common mild adverse events is a common feature of a drug.

Generally the safety concerns that remove drugs from the market or require label change are rare but medically very serious. Breaking it down in these frameworks gives less importance to the rare serious than the current approach of balancing it all in the mind in one go. Which is more appropriate? There is a tendency for people to overestimate how likely a rare event is and under estimate a common event, but at the same time there are issues with the way the weighting was done that could have underestimated the weight for PML.

The use of sensitivity analysis is key in a benefit-risk analysis. This shows which variables are driving the assessment, and focuses the discussion onto these specific weights and incidence measures. This demonstrates the value of information of these variables and may motivate more work to be done to better estimate them.

Stochastic sensitivity analysis was not performed in this case study. It could be argued that as the weights given to outcomes are decisions themselves it is inappropriate to assign stochastic variability to them. However, the outcome measures are estimated with uncertainty, and this uncertainty could be pushed through the benefit-risk model by, say, Monte Carlo simulation. This would lead to uncertainty in the benefit-risk balance. Ways of visualizing this uncertainty at the level of the individual outcomes, or benefit and risk categories, would also need to be considered.

Benefit-risk analysis employed by either of these methods is not an automated process where an algorithm is applied and an answer is generated. It is a framework to give clarity to decision makers. Expert opinion is central and many aspects are subjective, from the choice of outcomes to put in the value tree, to the selection of data and elicitation of weights. Benefit-risk analysis gives structure to enable discussion, ensures a wide breath of outcomes are considered and helps decision makes focus on the salient issues.

Section 8 References

(1) Rosati G. The prevalence of multiple sclerosis in the world: an update. Neurological sciences 2001;22(2):117-39.

(2) Compston A, Coles A. Multiple sclerosis. The Lancet 2002 Apr 6;359(9313):1221-31.

(3) GUIDELINE ON CLINICAL INVESTIGATION OF MEDICINAL PRODUCTS FOR THE TREATMENT OF MULTIPLE SCLEROSIS, CPMP/EWP/561/98 Rev. 1. http://www ema europa eu/ema/pages/includes/document/open_document jsp?webContentId=WC500003485 2006 November 16

(4) Assessment report for TYSABRI. EMA/H/C/000603/A20/0029 . http://www ema europa eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/000603/WC500095872 pdf 2010 May 10

(5) Kappos L, Bates D, Edan G, Eraksoy M, Garcia-Merino A, Grigoriadis N, et al. Natalizumab treatment for multiple sclerosis: updated recommendations for patient selection and monitoring. The Lancet Neurology 2011 Aug;10(8):745-58.

(6) Transatlantic workshop on Drug-related Progressive Multifocal Leukoencephalopathy (PML): Workshop Proceedings. EMA/674403/2011 Rev. 1. http://www ema europa eu/docs/en_GB/document_library/Report/2011/09/WC500111562 pdf 2011 September 8

(7) Third report on the progress of the interaction with patients' and consumers' organisations during 2009. EMA/MB/117170/2010. http://www ema europa eu/docs/en_GB/document_library/Report/2010/11/WC500099257 pdf 2010 October 7

(8) Coplan PM, Noel RA, Levitan BS, Ferguson J, Mussen F. Development of a framework for enhancing the transparency, reproducibility and communication of the benefit-risk balance of medicines. Clin Pharmacol Ther 2011 Feb;89(2):312-5.

(9) Levitan BS, Andrews EB, Gilsenan A, Ferguson J, Noel RA, Coplan PM, et al. Application of the BRAT framework to case studies: observations and insights. Clin Pharmacol Ther 2011 Feb;89(2):217-24.

(10) Holden WL, Juhaeri J, Dai W. Benefit-risk analysis: a proposal using quantitative methods. 2003.

(11) Holden WL, Juhaeri J, Dai W. Benefit-risk analysis: examples using quantitative methods. 2003.

(12) Sutton AJ, Cooper NJ, Abrams KR, Lambert PC, Jones DR. A Bayesian approach to evaluating net clinical benefit allowed for parameter uncertainty. 2005.

(13) John S.Hammond, Ralph L.Keeney, Howard Raïffa. Smart choices: a practical guide to making better decisions. Harvard Business Press; 1999.

Section 9 Appendix

Part 9.1 Data extraction

Table 21 literature searching and data selection strategy

<table>
<tr><td colspan="2">LITERATURE SEARCHING AND DATA SELECTION</td></tr>
<tr><td rowspan="2">Interferon Beta-1a
Interferon beta.m_titl. (title)
Beta interferon.m_titl.
Avonex.m_titl.
[1] OR [2] OR [3]
Multiple sclerosis.mp. and multiple sclerosis/ (Map term to subject heading)
[4] AND [5]</td><td>Glatiramer Acetate
Glatirameracetate.m_titl. (title)
Copaxone.m_titl.
Copolymer 1.m_titl.
Cop 1.m_titl.
[1] OR [2] OR [3] OR [4]</td></tr>
<tr><td>Natalizumab
Natalizumab.m_titl. (title)
Tysabri.m_titl.
[1] OR [2]</td></tr>
<tr><td colspan="2">Additional limits : English language AND Human AND (RCT or controlled clinical trial or phase II clinical trial or phase III clinical trial or phase IV clinical trial)</td></tr>
<tr><td colspan="2">Inclusion criteria for title and abstract screening: Clinical Trials for Relapsing-Remitting Multiple Sclerosis</td></tr>
<tr><td colspan="2">Exclusion criteria for title and abstract screening: Other Interferon beta not interferon beta-1a</td></tr>
<tr><td colspan="2">Inclusion Criteria for primary input data in MCDA model: Randomized placebo-controlled phase III clinical trial</td></tr>
</table>

Table 22 Data source table

GROUP	CATERGORY	OUTCOME	STATISTIC	SOURCE	DRUG	ACTIVE			PLACEBO			ACTIVE/PLACEBO		
Benefit						EST	LOW CI	UPP CI	EST	LOW CI	UPP CI	EST	LOW CI	UPP CI
	Relapse		Annualized relapse rate	Polman 2006 and EPAR	Natalizumab	0.23	0.19	0.28	0.73	0.62	0.87	0.32	0.26	0.4
				Jacobs 1996	interferon beta-1a	0.67			0.82			0.82	0.56	1.20
				Johnson 1998	Glatiramer acetate	0.65			0.91			0.71	0.47	1.08
	Disability progression		6-month confirmed % progressing after 2 years	EPAR	Natalizumab	11%			23%					
			6-month confirmed hazard	Polman 2006	Natalizumab	0.078			0.174			0.46	0.33	0.64
			6-month confirmed % progressing after 2 years	Jacobs 1996	interferon beta-1a	21.9%			34.9%					
			6-month confirmed hazard	Jacobs 1996	Interferon beta-1a	0.165			0.286			0.58	0.32	1.03
			3-month confirmed % progressing after 2 years	Johnson 1995	Glatiramer acetate	21.6%			24.6%					
			Ratio between 3-month and 6-month	Kappos 2010	Glatiramer acetate	0.71			0.79					
			6-month confirmed % progressing after 2 years		Glatiramer acetate	15.3%			19.4%					

			6-month confirmed hazard	Johnson 1995	Glatiramer acetate	0.111			0.144			0.77	0.41	1.46
		Convenience	Route and frequency of administration	Polman 2006	Natalizumab	iv qm hosp								
				Jacobs 1996	Interferon beta-1a	im qw								
				Johnson 1995	Glatiramer acetate	sc od								
Risk						ACTIVE			PLACEBO			RATIO		
						n	N	%	n	N	%			
	Infection	Reactivation of serious herpes viral infections		Polman 2006	Natalizumab	0	627	0	0	312	0%	1.00		
				Jacobs 1996	Interferon beta-1a	0	158	0	0	143	0%	1		
				Johnson 1995	Glatiramer acetate	0	125	0	0	126	0%	1		
		PML		Kappos 2011	Natalizumab	5	3722	0.151 0%	0	3722	0%			
				Jacobs 1996	Interferon beta-1a	0	158	0%	0	143	0%			
				Johnson 1995	Glatiramer acetate	0	125	0%	0	126	0%			
	Reproduction toxicity	Congential abnormalities		Polman 2006	Natalizumab	0	627	0	0	312	0			

				Jacobs 1996	Interferon beta-1a	0	158	0	0	143	0	
				Johnson 1995	Glatiramer acetate	0	125	0	0	126	0	
	Liver Toxicity	Transaminases elevation	ALT >5x ULN	Polman 2006	Natalizumab	31	627	5%	12	312	4%	1.25
				Jacobs 1996	Interferon beta-1a	0	158	0%	0	143	0%	1
				Johnson 1995	Glatiramer acetate	0	125	0%	0	126	0%	1
	Neurological	Seizures		Polman 2006	Natalizumab		627	0.50 %		312	0.50 %	
				FDA	Interferon beta-1a		158	3%		143	0%	
				Johnson 1995	Glatiramer acetate		125	0%		126	0%	
	Others	Infusion reactions/injection reactions		Polman 2006	Natalizumab	148	627	24%	55	312	18%	1.34
				FDA	Interferon beta-1a	20	158	13%	18	143	13%	1.00
				Johnson 1995	Glatiramer acetate		125	90%		126	59%	1.53
		Hypersensitivity Reactions		Polman 2006	Natalizumab	25	627	4%	0	312	0%	

				Jacobs 1996	Interferon beta-1a		158	0		143	0	1
				FDA	Glatiramer acetate	17	563	3%	11	564	2%	1.50
		Flu-like reactions		Polman 2006	Natalizumab		627	0%		312	0%	1
				Jacobs 1996	Interferon beta-1a	96	158	61%	57	143	40%	1.52
				Johnson 1995	Glatiramer acetate		125	0		126	0%	1

Part 9.2 Example questionnaire

RELAPSES AND DISABILITY PROGRESSION

Outcome level weights

The RRMS population is in a situation where they will experience:

Relapse: new or recurrent neurologic symptoms not associated with fever or infection that lasted for at least 24 hours and were accompanied by new neurologic signs found by the examining neurologist

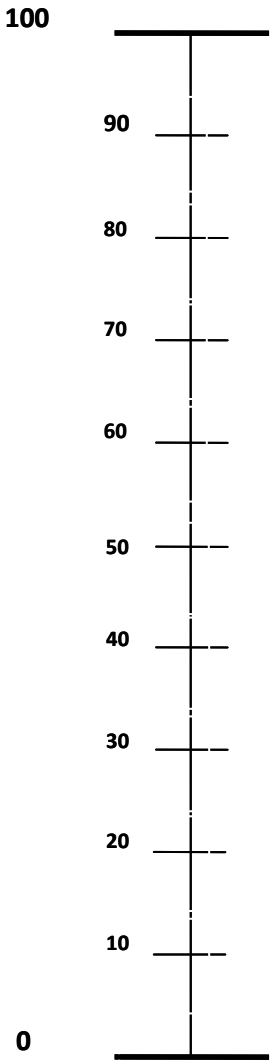
Disability progression: A >= 1.0 point increase in EDSS score with a baseline score >= 1.0 or a >=1.5 point increase in score with a baseline EDSS score of 0. Confirmed 6 months later.

You have the choice of changing the situation so they will avoid one of these events.

Rank the two outcomes in terms of how important it is to you to avoid each one.

Give the most important outcome a score of 100 and place the other outcome on the scale to reflect its relative importance.

Outcome	**Rank**
Avoid Relapse	
Avoid Disability Progression	



Most important outcome:

……

……

Part 9.3 Weights elicited

Table 23 Details of the weights elicited in the value tree hierarchy.

Level 1 Group	Outcome	Raw	Calibrated	Normalized	Level 2 Category	Outcome	Raw	Calibrated	Normalized	Level 3	Raw	Calibrated	Normalized	Weight	Adjusted	**Nomalized Final Weights**
Benefit	Convenience s.c. od --> Oral od	1	0.01	0.01	NA		100	1	1.00	Convenience s.c. od --> Oral od		1	1	1%	1%	**1%**
	Avoid disability progresison	10	0.17	0.09	NA		100	1.7	1.00	Avoid disability progresison	100	1	0.59	6%	6%	**5%**
										Avoid relapses	70	0.7	0.41	4%	8%	**8%**
Risk	Avoid PML	100	1.61	0.90	Infections	Avoid PML	100	1.12	0.70	Avoid PML	100	1	0.89	56%	56%	**54%**
										Avoid reactivation of herpes viral infection	12	0.12	0.11	7%	7%	**6%**
					Liver	Avoid transaminases elevation	20	0.2	0.12	Avoid transaminases elevation		1	1.00	11%	11%	**11%**
					Neurological disorders	Avoid seizures	10	0.1	0.06	Avoid seizures		1	1.00	6%	6%	**5%**
					Other	Avoid infusion/_njection site reactions	5	0.09	0.06	Aviod hypersensitivity reactions	40	0.4	0.22	1%	1%	**1%**
										Avoid infusion/_njection site reactions	100	1	0.56	3%	3%	**3%**
										Avoid flu-like	40	0.4	0.22	1%	1%	**1%**

		Reproductive	Congenital abnormalities	10	0.1	0.06	reations Congenital abnormalities	1	1.00	6%	6%	**5%**	

Part 9.4 People

Person	Areas of expertise
Richard Nixon	Statistics, Decision analysis
Thai Son Tong Nguyen	Pharmacology
Isabelle Stoeckert	Regulatory Affairs perspective, HC preferences
Christoph Dierig	Biostatistics – Integrated Analysis
Silvia Kuhls	Biostatistics – Integrated Analysis
Gemma Hodgson	Statistics
John Pears	Physician
Ioanna Tzoulaki	Statistics
Dolores Montero	Head pharmacovigilance at Agencia Española de Medicamentos y Productos Sanitarios