Australian Government
Department of Health
Therapeutic Goods Administration

AusPAR Attachment 3

Extract from the Supplementary Clinical Evaluation Report for Daclizumab

Proprietary Product Name: Zinbryta

Sponsor: Biogen Australia Pty Ltd

May 2017

TGA Health Safety Regulation

About the Therapeutic Goods Administration (TGA)

- The Therapeutic Goods Administration (TGA) is part of the Australian Government Department of Health, and is responsible for regulating medicines and medical devices.
- The TGA administers the *Therapeutic Goods Act 1989* (the Act), applying a risk management approach designed to ensure therapeutic goods supplied in Australia meet acceptable standards of quality, safety and efficacy (performance), when necessary.
- The work of the TGA is based on applying scientific and clinical expertise to decision-making, to ensure that the benefits to consumers outweigh any risks associated with the use of medicines and medical devices.
- The TGA relies on the public, healthcare professionals and industry to report problems with medicines or medical devices. TGA investigates reports received by it to determine any necessary regulatory action.
- To report a problem with a medicine or medical device, please see the information on the TGA website <https://www.tga.gov.au>.

About the Extract from the Clinical Evaluation Report

- This document provides a more detailed evaluation of the clinical findings, extracted from the Clinical Evaluation Report (CER) prepared by the TGA. This extract does not include sections from the CER regarding product documentation or post market activities.
- The words [Information redacted], where they appear in this document, indicate that confidential information has been deleted.
- For the most recent Product Information (PI), please refer to the TGA website <https://www.tga.gov.au/product-information-pi>.

Contents

List of common abbreviations

Abbreviation	Meaning
Ab	Antibody
AE	Adverse event
ANCOVA	Analysis of covariance
ARR	Annualised Relapse Rate
Avonex	IFN β-1a (tradename)
BBB	Blood-brain barrier
BG-12	Dimethyl fumarate
CD25	Interleukin-2 receptor, alpha subunit
CER	Clinical Evaluation Report
CI	Confidence interval
CNS	Central nervous system
DAC HYP	Daclizumab High Yield Process
DAC Penzberg	Daclizumab, Penzberg investigational form
EDSS	Expanded Disability Status Scale
EE	Efficacy-evaluable (population)
EEDS	Expanded Disability Status Scale
EMA	European Medicines Agency
Gd	Gadolinium
IFN	Interferon
IFN-β	Interferon beta
IgG	Immunoglobulin G
IL-2	Interleukin-2
IL-2Rα	Interleukin-2 receptor, alpha subunit
IM	Intramuscular

Abbreviation	Meaning
INEC	Independent Neurology Evaluation Committee
INEC	Independent Neurology Evaluation Committee
ISR	Injection site reaction
ITT	Intent-to-treat
IVRS	Interactive Voice Response System
JC	John Cunningham
LTi	Lymphoid tissue inducer (cells)
MAb	Monoclonal antibody
MRI	Magnetic Resonance Imaging
MS	Multiple sclerosis
MSFC	Multiple Sclerosis Functional Composite
MSIS-29	Multiple Sclerosis Impact Scale 29
NK	Natural killer (cells)
PD	Pharmacodynamic
PFS	Pre-filled syringe
PI	Product Information
PK	Pharmacokinetic
PPMS	Primary progressive multiple sclerosis
PRMS	Primary-relapsing multiple sclerosis
QoL	Quality of life
RRMS	Relapsing and remitting multiple sclerosis
SC	Subcutaneous(ly)
SCE	Summary of Clinical Efficacy
SCER	Supplementary Clinical Evaluation Report
SCS	Summary of Clinical Safety

Abbreviation	Meaning
SDMT	Symbol Digit Modalities Test
SOC	System organ class
SPMS	Secondary progressive multiple sclerosis
TGA	Therapeutic Goods Administration
VFT	Visual Function Test

1. Introduction

This submission proposes to register the new active substance daclizumab.

The submission proposes registration of the following dosage forms and strengths:

- Zinbryta (daclizumab) 150 mg/ml solution for injection Pre-filled Pen
- Zinbryta (daclizumab) 150 mg/ml solution for injection Pre-filled Syringe (PFS)

Daclizumab is a humanised immunoglobulin G (IgG) antibody (Ab), or more explicitly, a IgG1 monoclonal antibody (MAb) that binds specifically to the alpha subunit of the interleukin-2 receptor (IL-2Rα, also known as CD25) producing immunomodulatory effects by selectively blocking signalling through high affinity IL-2 receptors while leaving interleukin-2 (IL-2) signalling by intermediate affinity IL-2 receptors intact.

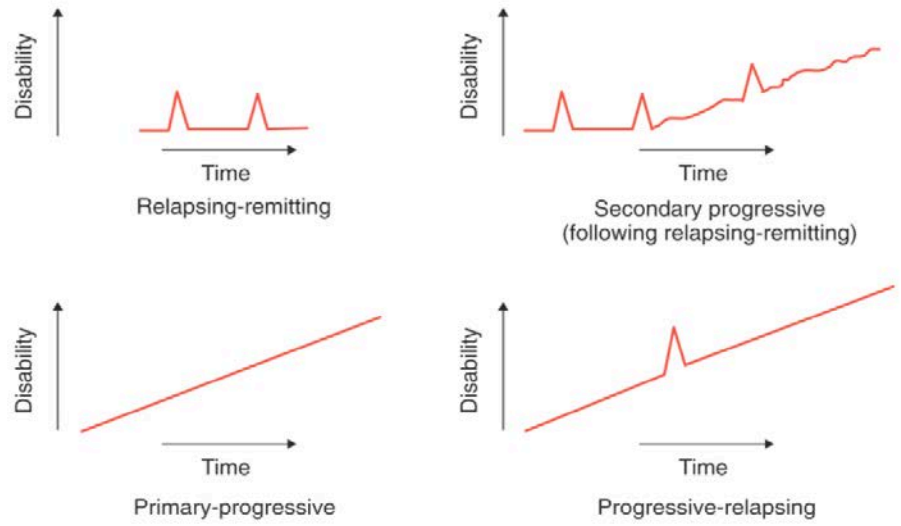
The sponsor's preferred indication, according to the proposed Product Information (PI), is as follows:

'Zinbryta is indicated for the treatment of relapsing forms of multiple sclerosis (MS).'

As will be discussed, the clinical evaluator believes that this indication is too broad, because it does not match the entry criteria for the pivotal studies. The indication should be reworded so that it explicitly applies to relapsing and remitting multiple sclerosis (RRMS), and not to primary or secondary progressive multiple sclerosis.

2. Clinical rationale

Figure 1. Different clinical courses of multiple sclerosis[1]



Multiple sclerosis (MS) is a complex neurological condition characterised by inflammation and demyelination in the central nervous system (CNS). Several subtypes are recognised, based on the temporal pattern of disability, as illustrated in Figure 1, above. The most common form is RRMS, characterised by acute episodes of focal inflammation, usually followed by recovery. This often transforms into secondary progressive multiple sclerosis (SPMS), in which progression of disability occurs between relapses, or in the absence of identifiable relapses. Primary

[1] Adapted from: Lublin F et al. Defining the clinical course of multiple sclerosis: results of an international survey. Neurology 1996; 46:907–911.

progressive multiple sclerosis (PPMS) also occurs, in which the disease is progressive from the onset. Some patients appear to have a progressive course from the onset, but also have superimposed relapses, a combination known as progressive-relapsing multiple sclerosis (PRMS).

MS is thought to be primarily an autoimmune disease, although there are also some neurodegenerative elements. Most approved therapies for MS are thought to reduce the incidence of relapses by modifying the immune system, and in some cases this has been shown to reduce the accumulation of disability.

Daclizumab High Yield Process (DAC HYP) is also thought to produce its benefits in MS by modifying the immune response and reducing CNS inflammation. Specifically, by blocking high-affinity IL-2 receptors, it produces the following immunomodulatory effects:

- Selective antagonism of activated T-cell responses.
- Expansion of immunoregulatory $CD56^{bright}$ natural killer (NK) cells which have been shown to selectively decrease activated T cells.
- Reduction in lymphoid tissue inducer (LTi) cells which are associated with cortical inflammation and demyelination.

3. Contents of the clinical dossier

3.1. Scope of the clinical dossier

This document (Attachment 3) is based upon the Supplementary Clinical Evaluation Report (SCER) written in response to the provision of new information by the sponsor. Accordingly, this extract from the SCER should not be considered a comprehensive assessment of all submitted clinical data. It should be read in conjunction with Attachment 2, the extract from the Clinical Evaluation Report (CER).

The primary material being evaluated in this SCER consists of the pivotal efficacy study reports, the sponsor's response to key Clinical Questions from the European Medicines Agency (EMA), and parts of the Summary of Clinical Safety (SCS) relevant to safety issues of particular interest.

The SCS has already been evaluated, as discussed in the First Round CER, and the main safety conclusions of that CER are summarised at the beginning of Section: Safety in this document.

One problem the current evaluator had with the SCS was that it did not directly include the tables and figures being discussed, so it was difficult to confirm claims made in the text of the SCS as they were being made. To a limited extent, this deficiency was offset by including hyperlinks to the relevant tables and figures, but this format increases the risk that discrepancies could be missed. The current clinical evaluator has not reassessed all aspects of the SCS to find such potential discrepancies.

In summary, the clinical material being evaluated is as follow:

- 2 pivotal efficacy and safety studies, Study 205MS201 and Study 205MS301
- Summary of Clinical Efficacy (SCE)
- Summary of Clinical Safety (SCS)
- Sponsor's responses to EMA questions (questions 70, 94 and 95).

3.2. Paediatric data

Please see Attachment 2, Extract of the CER.

3.3. Good clinical practice

Please see Attachment 2, Extract of the CER.

4. Pharmacokinetics

Pharmacokinetic data was not assessed by this clinical evaluator. Please see Attachment 2, Extract of the CER for a full evaluation.

5. Pharmacodynamics

Pharmacodynamic data was not assessed by this clinical evaluator. Please see Attachment 2, extract of the CER for a full evaluation.

6. Dosage selection for the pivotal studies

6.1. Dosage finding studies

Dose selection for the pivotal efficacy studies was based on studies in subjects with RRMS, performed with an investigational form of daclizumab (DAC Penzberg), manufactured using a different process and cell line.

According to the sponsor:

'Dose selection for Study 205MS201 was based on results of the Phase 2 Study DAC-1012, which evaluated 2 different dose regimens using a prior investigational form of daclizumab (DAC Penzberg): a high-dose 2 mg/kg subcutaneous (SC) every 2 weeks regimen (equivalent of 300 mg every 4 weeks for a 75 kg body weight) and the low-dose 1 mg/kg subcutaneously every 4 weeks regimen (equivalent of 75 mg every 4 weeks for a 75 kg body weight).'

'Compared with placebo the effect of DAC Penzberg on reducing new gadolinium (Gd) enhancing lesions, the primary endpoint of Study DAC-1012, was robust and statistically significant in the high-dose arm 2 mg/kg every 2 weeks (p = 0.0038), but was marginal and not statistically significant in the low-dose arm 1 mg/kg every 4 weeks (p = 0.5138). Safety was similar between the low-dose and high-dose regimens. Based on the results of Study DAC-1012, two DAC HYP dosing regimens (150 mg and 300 mg SC every 4 weeks) were selected for further evaluation in Study 205MS201 based on the following considerations:

- *The low-dose regimen from Study DAC-1012, which is approximately equivalent to a fixed-dose regimen of 75 mg SC every 4 weeks, was considered to be below the lowest efficacious dose. Furthermore, this regimen showed no evidence for an improved safety profile compared to the high-dose regimen. Therefore, DAC HYP doses that were expected to provide similar exposures were not evaluated further.*
- *DAC HYP 300 mg SC every 4 weeks was projected to be approximately equal to the highest efficacious dose (2 mg/kg SC every 2 weeks) evaluated in Study DAC-1012.*
- *DAC HYP 150 mg SC every 4 weeks was projected to be a lowest efficacious dose since it was between the low-dose and high-dose arms in Study DAC-1012.'*

6.2. Evaluator's overall conclusions on dose selection

A full evaluation of these claims is beyond the scope of this SCER, but the sponsor's selection of 150mg and 300mg as doses worthy of further study appears broadly reasonable.

Based on the results of DAC-1012, 150 mg SC 4-weekly and 300 mg SC 4-weekly were selected for the Phase 2 placebo-controlled study, 205MS201 (later designated as a pivotal study). In Study 205MS201, no difference in efficacy was observed between the 150 mg and 300 mg doses, so the lower dose was selected for the Phase 3 active-controlled study, 205MS301.

As noted by the First Round clinical evaluator, this development path suggests that efficacy plateaus above 150 mg, but it does not establish with certainty whether lower doses could still achieve comparable efficacy with an improved safety profile. The evidence suggests that, for DAC Penzberg, the optimal dose is greater than 75mg 4-weekly, and may be as high as 300 mg 4-weekly. To the extent that DAC HYP is equivalent to DAC Penzberg, Study 205MS201 further narrows down the optimal dose to somewhere above 75mg and up to 150 mg. It is not clear, though, that a 2-weekly dose with a different preparation of daclizumab is sufficient to guide dosing with a 4-weekly regimen of DAC HYP. Also, there is a two-fold range of doses between 75mg and 150 mg, leaving a wide range of doses untested. This represents a significant deficiency in the study program.

7. Clinical efficacy

7.1. Pivotal efficacy studies

7.1.1. Study 205MS201

Study 205MS201 was a *'multicentre, double-blind, placebo-controlled, dose-ranging study to determine the safety and efficacy of DAC HYP as a monotherapy treatment in subjects with RRMS.'*

7.1.1.1. Study design, objectives, locations and dates

Study 205MS201 (n = 600) was a placebo-controlled, double-blind, dose-ranging study in which two different doses of DAC HYP (the proposed dose of 150 mg, as well as 300 mg) were compared to placebo in subjects with RRMS. Per study report, this study was designated as a Phase 2 study, but was subsequently submitted as a pivotal study. It was only modest in size (600 subjects in the ITT population) and duration (52 weeks), so it would not be considered adequate as a standalone pivotal study. Also, it appears to have been designed with the expectation that 300 mg would be the dose subsequently taken to Phase 3 studies, which creates some difficulties in interpreting the statistical results in the 150 mg group. Despite these limitations, the study had most of the features required for a major MS efficacy study and it can be considered a pivotal study alongside the Phase 3 Study 205MS301 ('MS301'), which used an active control (IFN-beta-1A) in comparison to a single DAC HYP dose of 150 mg for up to 144 weeks.

The study ran from 15 February 2008 to 30 August 2011, and randomised a total of 621 subjects at 78 sites in 9 countries worldwide: the Czech Republic, Germany, Hungary, India, Poland, Russia, Turkey, the Ukraine, and the United Kingdom.

Inclusion criteria

The main inclusion criteria were:

- Male and female subjects between the ages of 18 and 55 years, inclusive

- A confirmed diagnosis of RRMS according to McDonald Criteria (numbers: 1 to 4)[2]
- Baseline Expanded Disability Status Scale (EDSS) between 0.0 and 5.0, inclusive
- Subjects had experienced at least 1 relapse within the 12 months prior to randomisation with a cranial magnetic resonance imaging (MRI) demonstrating lesions consistent with MS or had shown evidence of Gd-enhancing lesions of the brain on an MRI performed within the 6 weeks prior to randomisation.[3]

The main exclusion criterion was:

- Diagnosis of primary progressive, secondary progressive or progressive relapsing MS.

These criteria are reasonably standard in large MS efficacy studies. In combination, the inclusion and exclusion criteria attempted to define a cohort of subjects with active RRMS but no substantial ongoing disease progression. The requirement for either one relapse in the last 12 months or an active Gd lesion on MRI ensured that subjects with quiescent disease were not eligible. It is less clear that subjects with secondary progressive or progressive relapsing MS (SPMS or PRMS) were successfully excluded. It seems likely that many subjects with higher EDSS scores at baseline had, in part, reached a progressive phase of their illness, in which identifiable relapses were superimposed on a slowly progressive course. In practice, it is very difficult to distinguish the accumulation of disability that is due to incomplete recovery from relapses from disability that has increased between overt relapses, and the distinction is somewhat artificial, given that many radiological relapses are not recognise clinically. Notionally, SPMS and PRMS were listed as exclusion criteria, but the definitions of these categories required 3 months of continuous worsening, which may be very difficult to identify in clinical practice. Subjects with insidious progression might not get clearly worse over 3 months. Subjects with a fluctuating course including some accelerated periods of worsening around the time of their relapses could be classified by one neurologist as RRMS and by another as SPMS. Because of these ambiguities, subjects could have entered this study despite having SPMS or RPMS.

This is a problem faced by all major studies of RRMS, and the definitional approach taken in this study was acceptable. The same difficulty is faced by clinicians seeking to commence a disease-modifying agent. The potential inclusion of subjects with SPMS or PRMS is likely to have made it more difficult to demonstrate a treatment benefit, as these subjects are usually more treatment resistant.

A more serious concern is that the entry criteria do not match the target population identified in the sponsor's proposed PI. The proposed indication is:

'Zinbryta is indicated for the treatment of relapsing forms of multiple sclerosis (MS).'

PRMS is, by definition, a relapsing form of MS, and would be covered by the proposed indication even though it was an explicit exclusion criterion in this pivotal study. Similarly, most neurologists would consider a diagnosis of SPMS to be compatible with the occasional relapse, and therefore have a relapsing form of MS, which would be covered by the proposed indication. Furthermore, given that all SPMS begins with a relapsing course (by definition), SPMS could be considered a 'relapsing form' of MS even when the patient has reached a purely progressive

[2] McDonald criteria Numbers 1 to 4 were defined as: Criterion 1 = 2 or more relapses, 2 or more objective lesions. Criterion 2: 2 or more relapses, 1 objective lesion, and dissemination in space by MRI or positive CSF and 2 or more MRI lesions consistent with MS or further clinical attack involving different site. Criterion 3: 1 relapse, 2 or more objective lesions, and dissemination in time by MRI or second clinical attack. Criterion 4: 1 (mono-symptomatic) relapse, 1 objective lesion, dissemination in space by MRI or positive CSF and 2 or more MRI lesions consistent with MS, and dissemination in time by MRI or second clinical attack. Poleman (2005)

[3] According to the study report (Study 205MS201): "For inclusion purposes, a relapse was defined as neurologic signs and/or symptoms documented by a neurologist in the medical record and of at least 24 hours duration to be determined by the Investigator or the Treating Neurologist."

phase. The proposed indication in the PI should therefore be reworded to match the entry criteria of the pivotal studies.

Because of the broad range of EDSS scores permitted on study entry, it is also important to confirm whether the benefits were demonstrated across the full EDSS spectrum, making subgroup analyses particularly important.

Additional entry criteria were based on excluding: subjects with significant coexistent illnesses; those in whom exposure to an immune modifying agent could pose an unacceptable risk; and those in whom assessment of efficacy could be difficult because of use of other immune-modifying or disease-modifying agents.

7.1.1.2. Study treatments

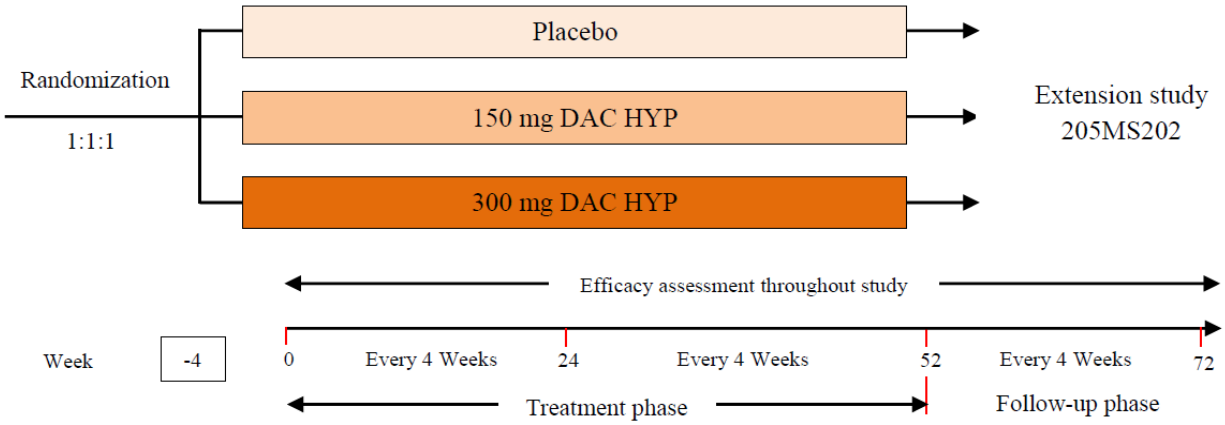
Subjects were randomised to one of three regimens in a 1:1:1 ratio:

- Placebo, administered by SC injection every 4 weeks
- DAC HYP 150 mg, administered by SC injection every 4 weeks
- DAC HYP 300 mg, administered by SC injection every 4 weeks

There was no dose titration phase. The planned duration of treatment was 52-weeks with an opportunity to enter a blinded extension study, Study 205MS202. The extension study is not described in this SCER.

Comment: For a description and evaluation of Study 205MS202 (the extension to Study 205MS201) please see Attachment 2, extract of the CER.

Figure 2. Study design, Study 205MS201



Rescue therapy with IFN-β was permitted for ethical reasons, to minimise the potential risks of untreated MS in the placebo arm. At the discretion of the treating clinician, IFN-β was used concomitantly with blinded study drug, starting at or after Month 6, provided the relapse had been confirmed by the Independent Neurology Evaluation Committee (INEC). Apart from rescue IFN-β, other disease-modifying agents were not allowed.

Intravenous methylprednisolone was allowed for treatment of relapses. All other systemic steroid therapy was prohibited.

Symptomatic therapy for spasticity, depression, or fatigue was allowed but clinicians were asked to optimise these as early as possible during screening in an attempt to maintain consistent treatments during the study.

7.1.1.3. Efficacy variables and outcomes

The main efficacy variables were:

- Brain MRI outcomes:

  - Total number of new Gd-enhancing lesions (Gd-enhancing lesions not present on MRI scan performed 4 weeks prior)
  - New or newly-enlarging T2 hyperintense lesions[4]
  - Volume of new T1 hypointense lesions
  - Volume of new or newly-enlarging T2 hyperintense lesions
  - Volume of non-Gd enhancing T1 hypointense ('blackholes') lesions
  - Brain atrophy
- Clinical outcomes:
  - Clinical relapses
  - EDSS
  - Subject Global Assessment (as measured by the Visual Analogue Scale)
  - Quality of life (QoL) questionnaires (EQ-5D, SF 12, and Multiple Sclerosis Impact Scale-29 (MSIS-29))
  - Relapses that were determined to meet protocol-defined criteria were subsequently evaluated by the Independent Neurology Evaluation Committee (INEC).

Primary efficacy outcome

The primary efficacy outcome was based on the Annualised Relapse Rate (ARR) between baseline and Week 52, calculated by dividing the number of relapses in each group by the total patient exposure in years.

This is an appropriate primary outcome. The ARR has been used in the majority of MS studies for decades. The primary endpoint used an adjusted form of the ARR, as is standard for MS studies of this nature, with statistical adjustments made on the basis of baseline prognostic factors (relapses, EDSS and age). As per study report:

'The primary analysis evaluated differences in the annualised relapse rate between each DAC HYP group versus placebo using a negative binomial regression model adjusting for the number of relapses in the year before study entry, baseline EDSS (EDSS ≤ 2.5 versus EDSS > 2.5 points), and baseline age (≤ 35 versus > 35 years).'

Secondary efficacy outcomes

Secondary efficacy outcomes included:

- The number of new Gd-enhancing lesions over 5 brain MRI scans at Weeks 8, 12, 16, 20, and 24 (calculated as the sum of these 5 MRIs) in a subset of subjects
- The number of new or newly-enlarging T2 hyperintense lesions at Week 52
- The proportion of relapsing subjects between baseline and Week 52
- QoL as measured by the MSIS-29 physical score at Week 52 compared to baseline.

These secondary endpoints were also reasonable. It is standard practice in MS studies to use radiological markers of disease activity as secondary endpoints. Radiological markers have the advantage of being objective and MRI is usually more sensitive than clinical relapse rate because many plaques may be clinically silent. On the other hand, a treatment that merely improved MRI markers without preventing neuronal dysfunction and disability would not be particularly

[4] T1 and T2 are MRI images are produced by dual spin echo frequency. Echo represents the signal received from the slice of interest in the body. T1 imaging is produced using short repetition and short echo times. T2 imaging is produced using long repletion and long echo times.

useful, so MRI markers are not suitable as primary endpoints. Gd lesions are indicators of local breakdown of the blood-brain barrier (BBB), which in the context of MS indicates probable inflammation. T2-weighted MRI is sensitive to water content, which increases in plaques and in other inflammatory areas, so enlarging or new T2 hyperintense lesions in an MS population are likely to indicate plaque growth; isolated lesions could be due to small vessel ischaemia, instead, but in an MS population a plaque is a more likely cause.

The proportion of relapsing subjects, which was used as a secondary endpoint, is also of interest, though it is strongly linked to ARR and generally does not provide major insights not already captured in ARR. Unlike ARR, this endpoint disregards second and subsequent in-study relapses from individuals, so it may be less sensitive to the inclusion of subjects with unusually aggressive disease and multiple relapses.

It is appropriate for a study in MS to use a measure of QoL and MSIS-29 is one validated tool suitable for this purpose. Unfortunately, this measure is subjective, and could potentially be affected by unblinding or other biases.

Other efficacy outcomes

Tertiary study objectives are listed below (as per study report) and included some safety assessments as well as efficacy measures:

'Tertiary objectives of this study were to determine:

- *The efficacy of DAC HYP in slowing the progression of disability as measured by at least a 1.0 point increase on the EDSS from baseline EDSS ≥1.0 that was sustained for 12 weeks, or at least a 1.5 point increase on the EDSS from baseline EDSS = 0 that was sustained for 12 weeks*
- *The efficacy of DAC HYP in reducing the number of new or newly-enlarging T2 hyperintense lesions at Week 24 compared to baseline*
- *The efficacy of DAC HYP in reducing the number of Gd-enhancing lesions at Week 52 compared to baseline*
- *The efficacy of DAC HYP in reducing the volume of new T1 hypointense lesions at Week 24 compared to baseline and at Week 52 compared to baseline*
- *The efficacy of DAC HYP in reducing the total lesion volume of new and newly enlarging T2 hyperintense lesions at Week 24 compared to baseline and at Week 52 compared to baseline*
- *The efficacy of DAC HYP in reducing the volume of non-Gd enhancing T1 hypointense ('blackholes') lesions at Week 24 compared to baseline and at Week 52 compared to baseline*
- *The efficacy of DAC HYP in reducing brain atrophy on MRI at Week 24 over the 52-week treatment period*
- *The efficacy of DAC HYP in reducing the total lesion volume of T2 hyperintense lesions over the 52-week treatment period*
- *The safety and tolerability of DAC HYP in subjects who have active, relapsing remitting forms of MS*
- *The efficacy of DAC HYP in slowing the time to relapse*
- *The efficacy of DAC HYP on slowing disability progression as measured by the change in EDSS from baseline to Week 52*
- *The efficacy of DAC HYP in improving the subject's global impression of well-being as measured by a Visual Analogue Scale*
- *The efficacy of DAC HYP in improving quality of life as measured by the MSIS-29 psychological scale, the SF-12 Health Survey (SF-12), and the EQ-5D Health Survey (EQ-5D).'*

Disability progression

Of note, none of the primary or secondary endpoints in Study 205MS201 included a measure of disability progression, which was instead listed as the first of 13 tertiary endpoints.

Disability progression was defined as a 1-point worsening of the EDSS, sustained for 12 weeks. (In the case of subjects with a baseline EDSS of 0, a 1.5 worsening was required, which helps to ensure that the disability is clinically significant). This is a standard definition, similar to many other MS studies, which have also defined disability as a sustained EDSS worsening.

The requirement for EDSS worsening to last 12 weeks has two main effects: it gives subjects 12 weeks to recover from a relapse, making it less likely that a relapse will be misinterpreted as disability progression; it also means that, in a 52-week study, subjects must exhibit worsening within 40 weeks of the start of the study. A requirement for longer periods of sustained worsening would be expected to be more specific for true progression, but the endpoint would be less sensitive because of the shorter time period available in which the disability would need to start in order to be counted. For a 52-week study, a 12-week period of sustained worsening is an appropriate compromise.

Guidelines for the conduct of MS studies strongly argue that an ideal MS treatment would be one that slowed disease progression. Ultimately, the accumulation of disability is a major concern for patients and their clinicians. Despite this, most MS treatments currently available reached the market on the strength of their ability to prevent relapses. Many treatments have since been shown to reduce disability progression, as well, but the benefits for this endpoint are less clear cut than the benefits on relapse rate. This partly reflects the fact that disability may be less responsive to immune-modifying treatment, but it may also reflect the fact that disability endpoints are less sensitive than relapse rates for purely methodological reasons. These reasons include the difficulties in distinguishing progression from relapses and the slow rate of progression relative to study duration. Despite its clinical importance, disability progression has often been treated as a minor endpoint in pivotal MS studies, as in this study.

7.1.1.4. Randomisation and blinding methods

Subjects were randomised to each of the three treatment arms in a 1:1:1 ratio, using an Interactive Voice Response System (IVRS).

Blinding was attempted by using identical appearing SC syringes in all three treatment arms, and by using randomisation codes that were not available to clinicians involved in the treatment and rating of patients.

It is possible that unblinding occurred because of tell-tale side effects, such as cutaneous reactions to DAC HYP.

The study took appropriate steps to separate treating and rating clinicians, as follows:

'To further protect the blind during the study, a separate treating neurologist and examining neurologist were designated at each investigational site. The treating neurologist functioned as the primary treating physician during the study. The examining neurologist conducted all EDSS evaluations and relapse assessments but was not involved in any other aspect of subject care and was instructed to limit all interactions with subjects to the minimum necessary to perform the required neurologic examinations. The examining neurologist remained blinded to treatment, AEs, concomitant medications, laboratory data, MRI data, and any other data that had the potential of revealing the treatment assignment.'

The sponsor apparently made no attempt to assess the degree of unblinding, which could have been achieved by asking subjects to guess their assigned treatment.

7.1.1.5. Analysis populations

The sponsor defined three study populations:

- Intent-to-treat (ITT) population: all randomised subjects who received at least 1 dose of study treatment. Note that subjects from one site were prospectively excluded because of systematic misdosing by the unblinded pharmacist.
- Efficacy-evaluable (EE) population: all subjects in the ITT population who had no missing MRI data from Weeks 8, 12, 16, 20, and 24 and did not take prohibited alternative MS medications.
- Safety population: all subjects who received at least 1 dose of study treatment and had at least 1 post-baseline assessment of the safety parameter being analysed.

The primary efficacy analysis was performed on the ITT population. The number of new Gd-enhancing lesions was evaluated using the EE population, and safety analyses were performed with the safety population.

7.1.1.6. Statistical methods

The primary endpoint was the difference in ARR between each active treatment and placebo. The primary analysis evaluated these differences with a negative binomial regression model, adjusting for the number of relapses in the year before study entry, the baseline EDSS (EDSS ≤ 2.5 versus EDSS > 2.5 points), and the baseline age (≤ 35 versus > 35 years). A traditional significance level of $p \leq 0.05$ was used.

The use of two active dose groups increases the chance of finding a significant difference relative to placebo in at least one active group. To control for this multiplicity, a sequential closed testing procedure was used to evaluate the dose groups. Statistical testing for efficacy endpoints used separate comparisons of the DAC HYP 300 mg group versus placebo and the DAC HYP 150 mg group versus placebo. Only if the comparison of 300 mg versus placebo was statistically significant ($p \leq 0.05$), was the comparison of 150 mg versus placebo to be tested. If the first comparison (300 mg) was not statistically significant, then the second comparison (150 mg) was not to be considered statistically significant, regardless of the p-value.

Secondary and other endpoints were summarised by treatment group, and tested for treatment differences by a number of different prospectively specified techniques:

- negative binomial regression (for number of new Gd+ lesions between Weeks 8 and 24, number of new or newly-enlarging T2 hyperintense lesions)
- a Cox-proportional hazards model (for time to first relapse, time to disability progression)
- an ordinal logistic regression model (for number of Gd+ lesions at Week 52)
- an analysis of covariance (ANCOVA) model (for change in EDSS, volume of new or enlarging T2 hyperintense lesions, volume of new T1 hypointense lesions, and QoL)
- Kaplan–Meier survival curve distribution (for the proportion of subjects who were relapse-free and the proportion of subjects who were progression-free).

To control for multiplicity of endpoints, secondary endpoints were rank prioritized, and if statistical significance was not achieved for an endpoint, endpoints of a lower rank were not considered significant.

Tertiary analyses including analysis of disability progression did not include any adjustments for multiplicity.

Overall, these analytical methods were broadly appropriate, but the tertiary endpoint analysis cannot be considered robust because of the lack of correction for multiplicity. Also, in reporting the benefits of active treatment on the proportion of subjects relapsing, the sponsor used an approach based directly on hazard ratios, with the result that the reported figures were inappropriately inflated. This issue is discussed in the results section for this study.

The statistical analysis plan was specified prospectively, but some additional analyses, which are a major focus of this supplementary evaluation report, were conducted in response to suggestions from the EMA, including a post-hoc analysis of the results according to baseline categorisation of subjects as having 'high disease activity' or 'low disease activity'.

Although these additional analyses were potentially informative, they cannot be considered statistically robust because of their post hoc nature.

7.1.1.7. Sample size

The sponsor justified the sample size as follows:

'It was assumed that if subjects were not allowed to add IFN-β during the study, the annualised relapse rate in the placebo group would be 0.50; however, because subjects were permitted to add IFN-β as a treatment for relapse, the annualised relapse rate in the placebo group would be reduced to 0.476 while the rate in the DAC HYP group would stay the same. In this setting, a sample size of 198 subjects per treatment group would have approximately 90% power to detect a 50% reduction in the annualized relapse rate between a DAC HYP treatment group and placebo. Power was estimated from simulations assuming a negative binomial distribution, a 10% drop out rate, and a 5% type 1 error rate. Based on these assumptions, a sample size of 594 subjects would be required for the study.'

These estimates appear reasonable, and the observed ARR in the placebo group (0.462) was very similar to the prospective estimate of 0.476. Furthermore, the study easily achieved statistical significance for its primary endpoint, indicating that it was, in fact, adequately powered for this endpoint.

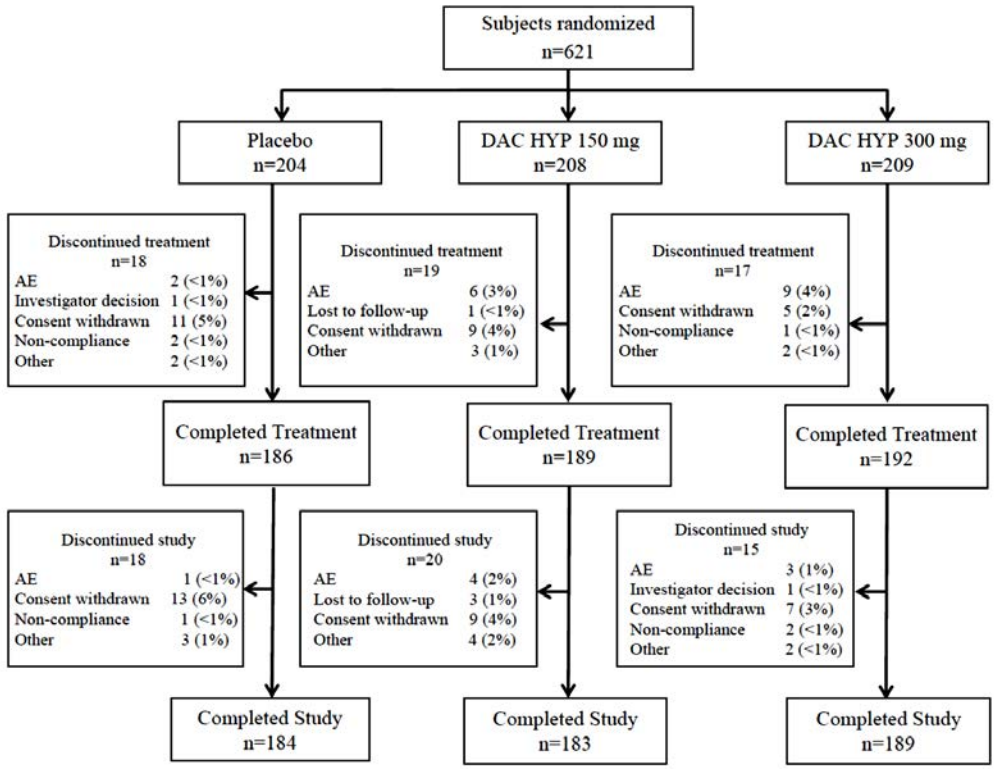
The study was not specifically powered for the tertiary endpoint of disability progression, and did not show a significant benefit for this endpoint.

7.1.1.8. Participant flow

A total of 621 subjects were randomised and all received study treatment. Due to systematic, non-random treatment at Site 93, a total of 21 subjects were excluded from the ITT population, leaving 196 subjects in the placebo group, 201 in the DAC HYP 150 mg group and 203 in the DAC HYP 300 mg group.

Of these, 186 completed placebo treatment, 189 completed DAC HYP 150 mg treatment, and 192 completed DAC HYP 300 mg treatment. A small number of subjects in each group completed treatment but did not complete the full follow-up period, as shown in the figure below; in some cases this reflects the enrolment of those subjects in the follow-up extension study.

Figure 3. Participant flow and subject disposition, Study 205MS201



Subjects who withdrew during the Study 201 follow-up period to enroll in the extension study were excluded from the total number of subjects who completed the study.

The reported completion rate constitutes fairly good follow-up for a complex study of this nature, and the withdrawals appear reasonably well-balanced across the three treatment groups. The most common reason for early discontinuation was withdrawal of consent, but adverse events (AEs) were more commonly listed as the reason for withdrawal in the active groups, which raises the possibility of withdrawal bias or unblinding.

7.1.1.9. Major protocol violations/deviations

The most serious protocol deviations occurred at Site 93, which was closed for study misconduct after it was discovered that the pharmacist dosed all 21 subjects with active DAC HYP rather than the randomised treatment assignments, including placebo. Data from these subjects were appropriately and prospectively excluded from the primary analysis. The sponsor also carried out sensitivity analyses that included the censored data, and this had no major effect on the results.

A clear summary table of all major protocol deviations was not provided, but has been requested as a clinical question to the sponsor.

7.1.1.10. Baseline data

Baseline demographics were similar across treatment groups. Baseline disease characteristics are summarised in Tables 1 to 4 below. The distribution of concomitant diseases, MS duration and the proportions of patients satisfying individual McDonald criteria were similar across groups. The relapse history was also similar across groups, including the number of relapses in the previous 3 years and previous 12 months, as well as the mean and median time since the last relapse before study entry.

Table 1. Medical history of subjects, Study 205MS201

Medical History

	Placebo	150 mg DAC HYP	300 mg DAC HYP	Total
Number of subjects randomized	204 (100)	208 (100)	209 (100)	621 (100)
Medical History				
Allergy	16 (8)	16 (8)	22 (11)	54 (9)
HEENT	34 (17)	37 (18)	36 (17)	107 (17)
Respiratory	9 (4)	14 (7)	14 (7)	37 (6)
Cardiovascular	34 (17)	34 (16)	39 (19)	107 (17)
Gastrointestinal	38 (19)	48 (23)	37 (18)	123 (20)
Hepatic	7 (3)	7 (3)	8 (4)	22 (4)
Genitourinary	47 (23)	38 (18)	38 (18)	123 (20)
Hematopoietic/Lymphatic	5 (2)	6 (3)	5 (2)	16 (3)
Neurological	204 (100)	208 (100)	209 (100)	621 (100)
Endocrine/Metabolic	19 (9)	16 (8)	24 (11)	59 (10)
Musculoskeletal	29 (14)	39 (19)	45 (22)	113 (18)
Dermatological	22 (11)	16 (8)	22 (11)	60 (10)
Psychosocial	26 (13)	19 (9)	23 (11)	68 (11)
Infectious Disease	18 (9)	14 (7)	17 (8)	49 (8)
Other	17 (8)	17 (8)	18 (9)	52 (8)

NOTE: Numbers in parentheses are percentages.

Table 2. MS history for subjects, Study 205MS201

History of Multiple Sclerosis

	Placebo	150 mg DAC HYP	300 mg DAC HYP	Total
Number of subjects randomized	204 (100)	208 (100)	209 (100)	621 (100)
Time since onset of the symptoms (Years)				
n	204	208	209	621
Mean	7.4	7.3	7.2	7.3
SD	6.93	6.30	6.38	6.53
Median	5.0	6.0	6.0	6.0
Min, Max	0, 30	0, 34	0, 34	0, 34
Time since diagnosis (Years)				
n	204	208	209	621
Mean	4.1	4.5	3.7	4.1
SD	5.26	4.96	4.00	4.77
Median	2.0	3.0	3.0	2.0
Min, Max	0, 26	0, 23	0, 21	0, 26
Dominant Hand				
Left	4 (2)	7 (3)	10 (5)	21 (3)
Right	200 (98)	201 (97)	199 (95)	600 (97)

NOTE: Numbers in parentheses are percentages.

Table 3. McDonald Criteria for subjects, Study 205MS201

Baseline McDonald Criteria Evaluation

	Placebo	150 mg DAC HYP	300 mg DAC HYP	Total
Number of subjects randomized	204 (100)	208 (100)	209 (100)	621 (100)
McDonald Criteria				
1 (a)	156 (76)	165 (79)	164 (78)	485 (78)
2 (b)	32 (16)	27 (13)	30 (14)	89 (14)
3 (c)	14 (7)	12 (6)	14 (7)	40 (6)
4 (d)	2 (<1)	4 (2)	1 (<1)	7 (1)

NOTE: Numbers in parentheses are percentages.
(a) 2 or more relapses, 2 or more objective lesions.
(b) 2 or more relapses, 1 objective lesion, and dissemination in space by MRI or positive CSF and 2 or more MRI lesions consistent with MS or further clinical attack involving different site.
(c) 1 relapse, 2 or more objective lesions, and dissemination in time by MRI or second clinical attack.
(d) 1 (mono-symptomatic) relapse, 1 objective lesion, dissemination in space by MRI or positive CSF and 2 or more MRI lesions consistent with MS, and dissemination in time by MRI or second clinical attack.

Table 4. Relapse history of subjects, Study 205MS201

	Placebo	150 mg DAC HYP	300 mg DAC HYP	Total
Number of subjects randomized	204 (100)	208 (100)	209 (100)	621 (100)
Number of relapse during the previous 3 years				
0	0	1 (<1)	0	1 (<1)
1	43 (21)	46 (22)	53 (25)	142 (23)
2	89 (44)	76 (37)	84 (40)	249 (40)
3	46 (23)	51 (25)	34 (16)	131 (21)
>=4	26 (13)	34 (16)	38 (18)	98 (16)
n	204	208	209	621
Mean	2.3	2.5	2.4	2.4
SD	1.14	1.29	1.52	1.32
Median	2.0	2.0	2.0	2.0
Min, Max	1, 7	0, 8	1, 12	0, 12

	Placebo	150 mg DAC HYP	300 mg DAC HYP	Total
Number of relapse during the past 12 months				
0	4 (2)	7 (3)	1 (<1)	12 (2)
1	136 (67)	129 (62)	150 (72)	415 (67)
2	56 (27)	62 (30)	46 (22)	164 (26)
3	7 (3)	8 (4)	10 (5)	25 (4)
>=4	1 (<1)	2 (<1)	2 (<1)	5 (<1)
n	204	208	209	621
Mean	1.3	1.4	1.3	1.4
SD	0.60	0.73	0.68	0.67
Median	1.0	1.0	1.0	1.0
Min, Max	0, 4	0, 6	0, 6	0, 6
Time since the most recent pre-study relapse in months (a)				
n	204	208	209	621
Mean	5.3	5.5	5.6	5.5
SD	3.27	3.33	2.94	3.18
Median	5.0	4.5	5.0	5.0
Min, Max	1, 24	1, 20	1, 14	1, 24

NOTE: Numbers in parentheses are percentages.
(a) Time is from the date of the most recent pre-study relapse to the date of randomization.

For radiological markers of disease severity, there was a slight imbalance across groups, suggesting more active disease in the 150 mg group compared to the placebo and 300 mg groups.

Across the entire study population, the average number of Gd lesions on the baseline MRI was 1.8 ± 3.78, and 44% of subjects had ≥ 1 Gd lesion. The median volume of T2 lesions was 4563.7 mm^3. The proportion of subjects with 1 Gd-enhancing lesion on the baseline MRI was higher than average in the DAC HYP 150 mg group: 52%, compared to 45% and 37% in the placebo and DAC HYP 300 mg groups, respectively. Similarly, the median volume of T2 hyperintense lesions was higher in the 150mg group: 5392 mm^3 in the DAC HYP 150 mg group, compared to 4492 mm^3 and 4113 mm^3 in the placebo and DAC HYP 300 mg groups, respectively.

These values were misquoted in the study report, as follows:

'The median volume of T2 hyperintense lesions was 5392 mm^3 in the DAC HYP 150 mg group compared to 4492 mm^3 in the DAC HYP 150 mg (sic) group and 4113 mm^3 in the placebo (sic) group.'

Table 5. Volume of T2 hyperintense lesions (mm^3) at baseline, Study 205MS201

	Placebo	150 mg DAC HYP	300 mg DAC HYP	Total
Volume of T2 hyperintense lesions (mm^3)				
0 (Q1)	3 (1)	1 (<1)	3 (1)	7 (1)
>0 - <1574 (Q1)	52 (25)	42 (20)	54 (26)	148 (24)
1574 - <4564 (Q2)	49 (24)	51 (25)	53 (25)	153 (25)
4564 - <11980 (Q3)	46 (23)	53 (25)	54 (26)	153 (25)
11980 - 52773 (Q4)	53 (26)	59 (28)	42 (20)	154 (25)
Missing	1 (<1)	2 (<1)	3 (1)	6 (<1)
n	203	206	206	615
Mean	8080.3	9888.9	7001.7	8324.8
SD	9421.62	11100.56	8076.42	9672.16
Median	4492.2	5392.1	4113.5	4563.7
Min, Max	0, 52773	0, 50972	0, 37823	0, 52773

NOTE: Numbers in parentheses are percentages.

Because they had more active scans, the 150 mg group might be expected to have more relapses during the study, which could potentially bias the study slightly against the 150 mg group for the primary endpoint of ARR. On the other hand, because a high proportion of subjects were having a radiological relapse at baseline, some clinical improvement in EDSS could be expected to arise purely from recovery from baseline relapses; this effect could disguise some progression; potentially this effect could have been more prominent in the 150 mg group. Overall, the discrepancy between groups for this baseline measure was minor and the groups were well-matched for clinical relapse history, so it is not expected to have modified the findings significantly. Also, post hoc comparisons of disability progression for subjects with high and low baseline disease activity did not find a major difference in the estimates of the treatment effect.

Overall, the treatment groups were adequately balanced and they were representative of the population in which DAC HYP would be used.

7.1.1.11. Results for the primary efficacy outcome

The ARR at 52 weeks was significantly lower for subjects randomised to active treatment, relative to the ARR observed with placebo. The adjusted ARR was 0.458 relapses/year in the placebo group, compared to 0.211 in the DAC HYP 150 mg group (a 54% relative reduction; 95% confidence interval (CI), 33% to 68%; $p < 0.0001$), and 0.230 in the DAC HYP 300 mg group (a 50% relative reduction; 95% CI, 28% to 65%; $p = 0.0002$). These results are summarised in the table below, reproduced from the sponsor's submission.

The meaning of the p-values cited next to footnote 'b' in the sponsor's table (see Table 6) was not clear, and the sponsor should be asked to clarify this.

Table 6. Annualised relapse rate by study treatment, Study 205MS201

Primary analysis – Annualized relapse rate between baseline and week 52 – Negative binomial regression
Page 1 of 2

	Placebo	150 mg DAC HYP	300 mg DAC HYP
Number of subjects in ITT population	196 (100)	201 (100)	203 (100)
Number of relapses			
0	127 (65)	163 (81)	163 (80)
1	52 (27)	33 (16)	34 (17)
2	15 (8)	5 (2)	5 (2)
3	2 (1)	0	1 (<1)
>= 4	0	0	0
Total number of relapses	88	43	47
Total subject-years followed	190.39	193.90	197.51
Unadjusted annualized relapse rate (a)	0.462	0.222	0.238
Adjusted relapse rate (95% CI) (b)	0.458 (0.370,0.566)	0.211 (0.155,0.287)	0.230 (0.172,0.308)
Rate ratio (95% CI) (b)		0.461 (0.318,0.668)	0.503 (0.352,0.721)
p-value vs placebo		<0.0001	0.0002
Subject relapse rate (c)			
n	196	201	203
Mean	0.484	0.229	0.250
SD	0.7958	0.5419	0.6024
Median	0.000	0.000	0.000
25th, 75th percentile	0.000, 1.001	0.000, 0.000	0.000, 0.000
Min, Max	0.00, 4.96	0.00, 4.20	0.00, 4.94

Note 1: Numbers in parentheses are percentages.
2: Data after subjects switched to alternative MS medications are excluded.
(a) Total number of relapses that occurred during the study divided by the total number of subject-years followed in the study.
(b) Estimated from a negative binomial regression model adjusted for the number of relapses in the 1 year prior to study entry (p= 0.005), baseline EDSS (<= 2.5 vs > 2.5, p= 0.411), and age (<= 35 vs > 35, p= 0.063).
(c) Number of relapses for each subject divided by the number of years followed in the study for that subject. Summary statistics are presented.

The observed reduction in relapse rate (approximately 50 to 54%, depending on which dose group is considered) is clinically meaningful and resembles the reported reductions in ARR observed in other recent pivotal studies of different disease-modifying agents, including fingolimod and dimethyl fumarate (BG-12). For the pivotal placebo-controlled fingolimod trial, the ARR was 0.18 in the active group, compared to 0.40 in the placebo group, a relative reduction of 55% ($p < 0.001$). In the pivotal study of dimethyl fumarate, the reduction in ARR was also similar:

'The annualised relapse rate at 2 years was 0.17 in the twice-daily BG-12 group and 0.19 in the thrice-daily BG-12 group, as compared with 0.36 in the placebo group, representing relative reductions of 53% and 48% with the two BG-12 regimens, respectively (p < 0.001 for the comparison of each BG-12 regimen with placebo).'[5]

Although comparisons across studies are not formally valid, these relatively recent studies had broadly similar entry criteria and definitions of relapse rate, as well as a similar relapse rate in their respective placebo groups, so they provide useful context for the findings in Study 205MS201.

The first disease-modifying treatments to be marketed for MS, including IFN-β and glatiramer acetate, were associated with an apparent reduction in relapse rate of approximately 30%, but those earlier studies may have recruited a more advanced cohort of MS patients. There has been a general trend to earlier treatment and to more favourable results in recent MS studies,

[5] Gold R et al; DEFINE Study Investigators. Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis. N Engl J Med. 2012 Sep 20;367(12):1098-107.

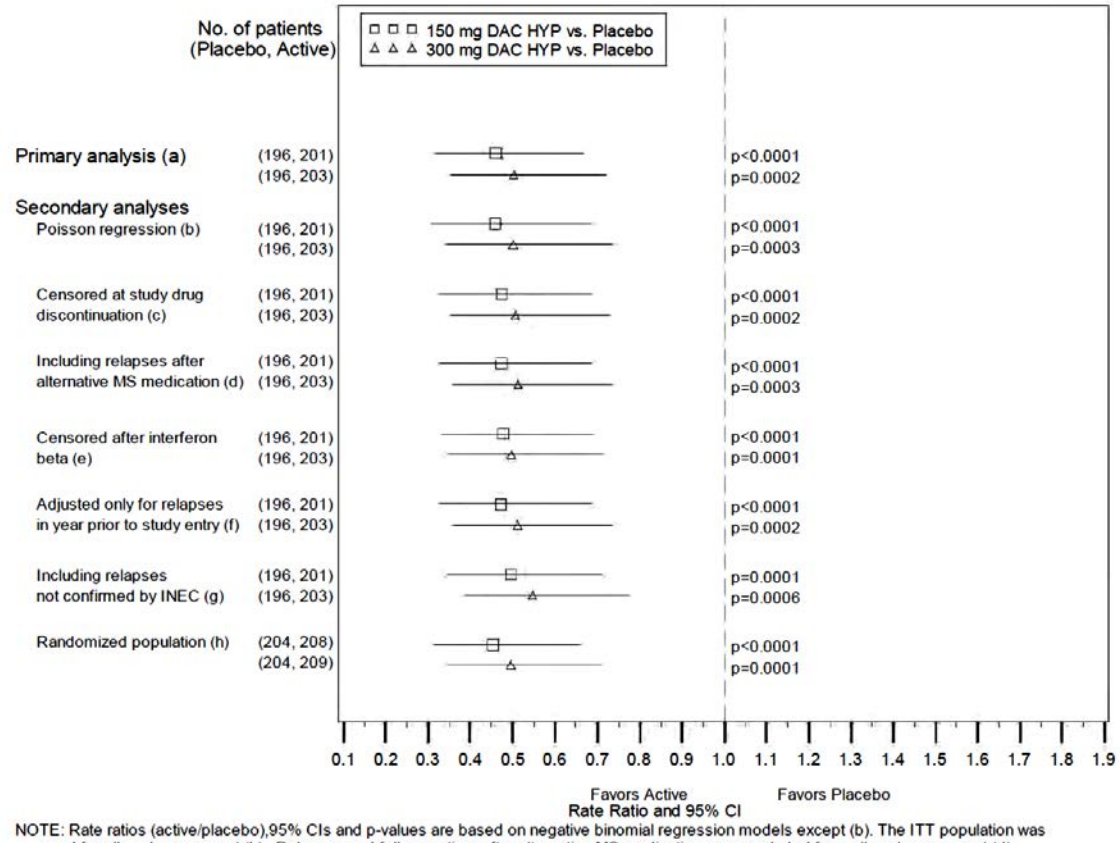
compared to the initial studies performed in the 1990s, so comparison across treatment eras is unreliable.

Given that individual relapses are themselves unpleasant and often disabling, patients would be expected to welcome any treatment that reduced relapses by approximately 50 to 54%. As some accumulation of disability in MS is directly related to incomplete recovery from individual relapses, the prevention of half the expected number of relapses would be expected to reduce long-term disability.

Unfortunately, as discussed elsewhere in this report, the extent to which reductions in ARR correlate with improvements in disease progression has been generally disappointing in MS research. It is therefore necessary to demonstrate such benefits directly. In this study, as will be discussed, a direct benefit in terms of EDSS progression was only confirmed for some progression-related endpoints, but not for the main prospectively identified progression endpoint.

The sponsor also performed a number of additional analyses of the primary efficacy variable, using a variety of statistical techniques, with and without censoring of relapses after rescue therapy, and including unconfirmed relapses. The sponsor performed an analysis in which they only adjusted the ARR for relapses at baseline, instead of relapses, age and EDSS. As shown in the figure below, all of these sensitivity analyses produced concordant results, suggesting that the results of the primary analysis are robust, and did not depend strongly on arbitrary methodological choices.

Figure 4. Primary and additional analyses of annualised relapse rate, Study 205MS201



NOTE: Rate ratios (active/placebo),95% CIs and p-values are based on negative binomial regression models except (b). The ITT population was used for all analyses except (h). Relapses and follow up time after alternative MS medication were excluded from all analyses except (d). Relapses were required to be confirmed by INEC for all analyses except (g). Models (a), (b) and (h) were adjusted for the number of relapses in the 1 year prior to study entry plus baseline EDSS (<=2.5 vs >2.5) and age (<=35 vs >35). All other models were adjusted only for the number of relapses in the 1 year prior to study entry.

Overall, these primary efficacy results appear statistically and clinically robust. There was a slight trend in favour of the 150 mg dose over the 300 mg dose, but no formal dose comparison was attempted and the study was not powered to compare the two active doses. The clear

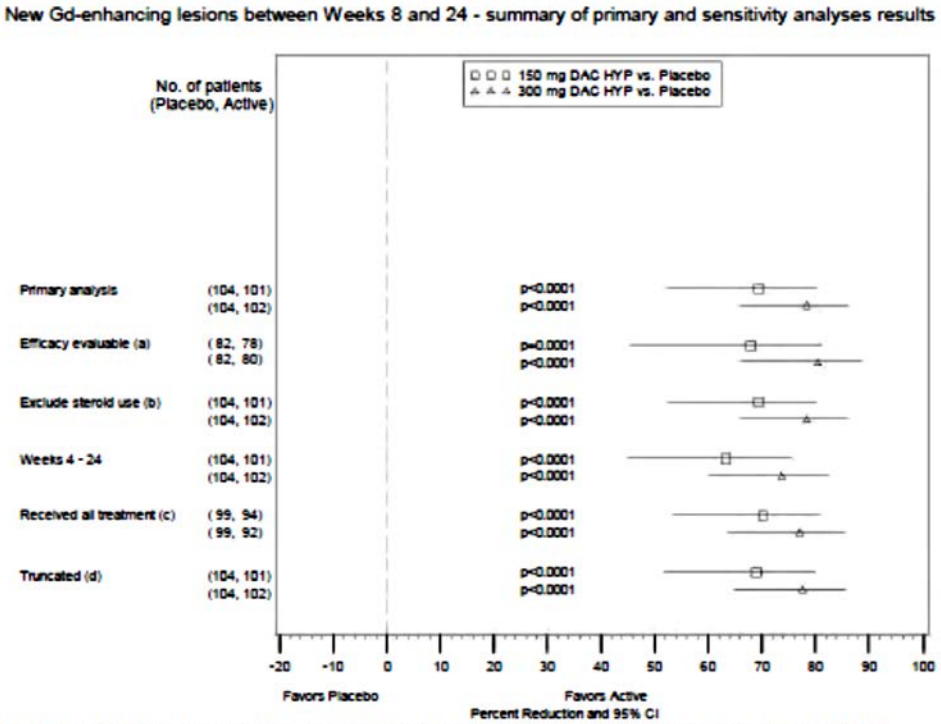
superiority of both active doses over placebo and the lack of any substantial differences in the 150 mg and 300 mg groups broadly justify the sponsor's subsequent development of 150 mg in favour of 300 mg, but leave open the possibility that lower doses could have had comparable efficacy and improve safety.

7.1.1.12. Results for other efficacy outcomes

Radiological measures

Gd-enhanced MRI scans were obtained at Weeks 8, 12, 16, 20, and 24 in a subset of subjects. With active treatment, the number of new Gd-enhancing lesions (calculated as the sum of new lesions across these 5 MRIs) was reduced versus placebo. DAC HYP 150 mg and 300 mg reduced the number of new Gd-enhancing lesions by 69% ($p < 0.0001$) and 78% ($p < 0.0001$) respectively, compared to placebo. This result appeared to be robust in a variety of sensitivity analyses, as shown in the Forest plot (Figure 5) below.

Figure 5. New Gd-enhancing lesions, primary and sensitivity analyses, Study 205MS201



NOTE: Percent reduction (active compared to placebo), 95% CIs and p-values are based on negative binomial regression models adjusted for the baseline number of Gd-enhancing lesions. The MRI intensive population was used for all analyses except (a). For subjects with missing data the last valid non baseline measurement was carried forward if the subject was missing only 1 or 2 consecutive post-baseline scans. Otherwise the mean based on treatment group and visit was used as the imputed value. MRIs taken after alternative MS medication are excluded and the imputed value is used instead.

(a) Analysis on the efficacy evaluable population. (b) Scans obtained within 24 days of steroid treatment are treated as missing. (c) Analysis on the subset of subjects who received all 6 assigned study drug administrations through Week 20. (d) Number of lesions truncated at 30.

New or newly-enlarging T2 hyperintense lesions at Week 52

The MRI data also suggested a substantial reduction in the development of T2 lesions. The adjusted mean number of new or newly enlarging T2 hyperintense lesions at Week 52 was 8.13 (95% CI: 6.65 to 9.94) in the placebo group, compared to 2.42 (95% CI: 1.96 to 2.99) in the DAC HYP 150 mg group ($p < 0.0001$) and 1.73 (95% CI: 1.39 to 2.15) in the DAC HYP 300 mg group ($p < 0.0001$). The relative reduction in the number of new or newly enlarging T2 lesions, compared to placebo, was 70% for DAC HYP 150 mg ($p < 0.0001$) and 79% for DAC HYP 300 mg ($p < 0.0001$), respectively.

Proportion of relapsing subjects

The Kaplan-Meier estimate for the proportion of subjects who relapsed by Week 52 was 36% in the placebo group, compared to 19% in the DAC HYP 150 mg and 20% in the DAC HYP 300 mg

group. The hazard ratio was 0.45 (95% CI: 0.30 to 0.67; p < 0.0001) in the DAC HYP 150 mg group compared to placebo and 0.49 (95% CI: 0.33 to 0.72; p < 0.0003) in the DAC HYP 300 mg group compared to placebo.

The sponsor and the previous CER have suggested that these results indicate that the proportion of relapsing subjects was reduced by 55% in the DAC HYP 150 mg group and by 51% in the DAC HYP 300 mg group, compared to placebo. This does not represent an accurate description of the results. The 55% and 51% reductions appear to have been derived directly from the hazard ratios of 0.45 and 0.49, respectively, and thus refer to the reduction in the proportion relapsing from the 'at-risk' (not-yet-relapsed) group at any one moment in time, but do not apply to the cohort over the 52-week time period as a whole. The cited reductions are not plausible. The proportion relapsing in the placebo group was 36%, so if the relative reduction in the proportion relapsing on active treatment was 50%, then 18% of those on active treatment (half of 36%) would have relapsed. Instead, more than 18% of subjects relapsed in both active groups (19% and 20% in the 150 mg and 300 mg groups, respectively), so the proportion relapsing cannot have been reduced by more than 50%.

The relative reduction in the proportion relapsing was actually $1 - {}^{19}/_{36}$ or 0.47 (47%) for the 150 mg group, and $1 - {}^{20}/_{36}$ or 0.44 (44%) in the 300 mg group. In other words, for the 150 mg group, the risk of relapse was 0.53 times the risk with placebo (0.5278 x 36% = 19%), and the risk of relapse with 300 mg was 0.56 times the risk with placebo (0.5556 x 36% = 20%). Slightly different values might be obtained if more significant figures were used for the initial proportions. The Study synopsis uses a value of 35% for the proportion relapsing in the placebo group, instead of 36% (possibly due to rounding errors), and if this value were used it would give even lower estimates of the relative reduction with active treatment.

The inflated estimates of risk reduction arise from disregarding the difference between instantaneous hazard ratios and overall risks for a cohort over an extended period of time. The result of conflating these two measures of risk is that the estimates of treatment benefits are exaggerated. By definition, instantaneous hazard ratios ignore subjects who have already experienced the hazard event, and are therefore based on a shrinking denominator, but clinical intuition and decision-making are based on the overall proportion of subjects experiencing the hazard event in a clinically meaningful time period, such as a year of treatment, and use the entire cohort as the denominator.

A review of the documents provided with the submission shows that this error is repeated throughout the submission, and it has been subsequently accepted in good faith by the First Round clinical evaluator. For instance, the Study Report cites the inflated estimates of the reduction in proportion of subjects relapsed (55% for the 150 mg group, instead of 47%; and 51% for the 300 mg group, instead of 44%). The inflated values are also used in the proposed PI, where it is stated that the relative risk reduction for the proportion relapsing on 150 mg was 55%. This should be corrected.

Quality of life

QoL was measured with the MSIS-29 physical score. MSIS-29 scores at Week 52 were compared to baseline. Results for this endpoint were not significant. A p-value that was nominally within the significance range was obtained for the dose group of secondary interest (150 mg) but not for the protocol-specified dose group of primary interest (300 mg). By the prospectively declared sequential closed testing procedure, significance for the 300 mg dose group had to achieve statistical significance before the 150 mg dose group could be tested.

Mean changes in MSIS-29 scores were small compared to the variability within each group, so it is difficult to draw any strong inferences and it is unclear whether the differences observed were clinically meaningful. The mean (± SD) change in the MSIS-29 physical score from baseline to Week 52 was 3.0 (± 13.52) in the placebo group, –1.0 (± 11.80) in the DAC HYP 150 mg group

(p = 0.0008 versus placebo), and 1.4 (± 13.53) in the DAC HYP 300 mg group (p = 0.1284 versus placebo).

Even the direction (sign) of the change in MSIS-29 was different across active groups and there was no consistent dose trend.

Overall, this endpoint provides no convincing support for the efficacy of DAC HYP. The sponsor claims that DAC HYP produces benefits in MSIS-29 scores, but this claim is not justified. Where the PI mentions the p-value of 0.0008, it should also mention that this was not formally significant.

Key tertiary endpoint Disability progression

Disability progression, a key tertiary endpoint, was generally underpowered because few subjects demonstrated confirmed disability progression within the time window required and the study was generally too short to provide robust estimates for this endpoint. The overall trends, however, were favourable, and the p-value for the 150 mg group was nominally in the appropriate range (≤ 0.05); this does not imply statistical significance, because the pre-specified approach for handling multiplicity was to use a sequential closed testing procedure, but it is a least suggestive of a benefit on disability progression.

The proportion of subjects with 12-week confirmed disability progression was 25 (13%) in the placebo group, 11 (5%) in the DAC HYP 150 mg group, and 15 (7%) in the DAC HYP 300 mg group. (In the adjusted estimate, the rates were 13.3%, 5.9% and 7.8%, respectively). Relative to placebo, the hazard ratio for disability progression was 0.43 (95% CI: 0.21 to 0.88) in the DAC HYP 150 mg group and 0.57 (95% CI: 0.30 to 1.09) in the DAC HYP 300 mg group.

Table 7. Time to sustained progression, Study 205MS201

Summary of time to sustained progression of disability measured by increase in EDSS

	Placebo	150 mg DAC HYP	300 mg DAC HYP
Number of subjects in ITT population	196 (100)	201 (100)	203 (100)
Number of subjects who progressed	25 (13)	11 (5)	15 (7)
Time (wk) to progression (a)			
25th percentile	NA	NA	NA
50th percentile	NA	NA	NA
Estimated proportion of subjects with progression at 52 weeks (a)	0.133	0.059	0.078
Hazard ratio and 95% CI (b)		0.43 (0.21-0.88)	0.57 (0.30-1.09)
p-value vs placebo (b)		0.0211	0.0905

NOTE: Sustained progression of disability is defined as at least a 1.0 point increase on the EDSS from a baseline EDSS >=1.0 sustained for 12 weeks or at least a 1.5 point increase on the EDSS from a baseline EDSS of 0 sustained for 12 weeks.
(a) Estimated time to progression and proportion of subjects with progression based on the Kaplan-Meier product limit method.
(b) Hazard ratio and p-value assessing the difference between the treatment groups were estimated from a Cox proportional hazards model. Covariates included were baseline EDSS (<=2.5 versus >2.5, p= 0.085), and age (<=35 versus >35, p= 0.172).

As with the proportion relapsing, discussed above, the sponsor and the First Round clinical evaluator cited reductions in the proportions progressing that were directly based on hazard ratios, and this is potentially misleading. The sponsor suggested that the risk of disability progression was reduced by 57% in the DAC HYP 150 mg group (p = 0.0211) and by 43% in the DAC HYP 300 mg group (p = 0.0905), compared with placebo. These calculations were based on hazard ratios and do not accurately reflect the reduction achieved by the entire cohort over one year, but the inflation effect is relatively minor because so few subjects reached the hazard of interest.

As a proportion of the placebo progression rate, the proportion progressing in the 150 mg dose group was 5.9/13.3 (44.4%), consistent with a reduction of 55.6% (not 57%, as claimed). For the 300 mg dose group, the relative proportion progressing was 7.8/13.3 (58.6%), consistent

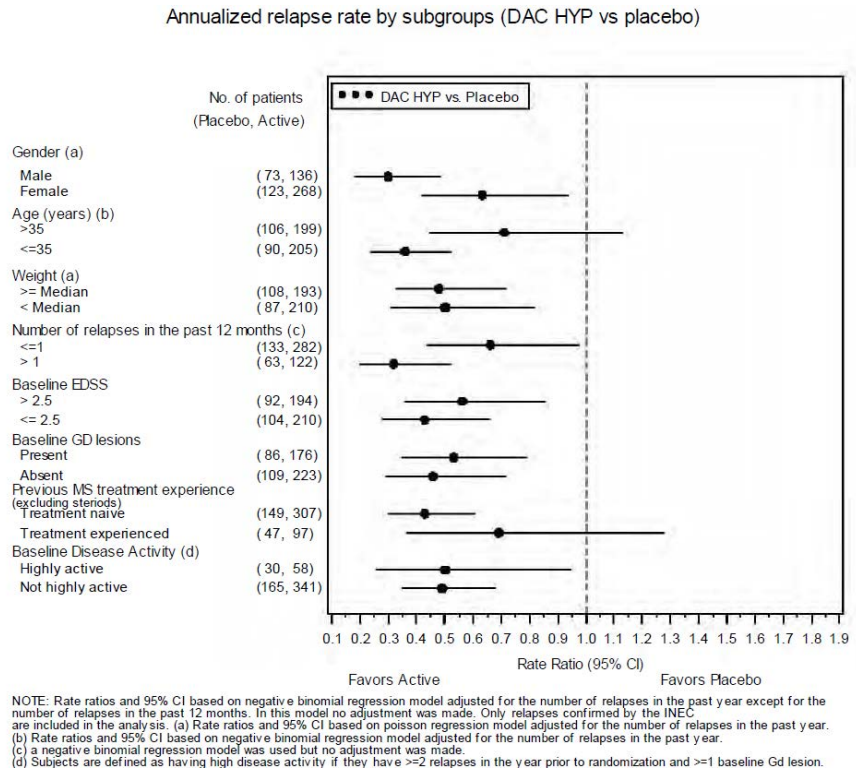
with a reduction of 41.4% (not 43%). These results are similar to those cited by the sponsor, but the PI should be changed to reflect the actual reduction in the proportion, not the inflated estimate based on hazard ratios. Also, a p-value of 0.0211 is cited in the proposed PI without any indication that this result was not statistically significant; the PI should be amended to indicate a non-significant result.

An additional minor endpoint consisted of the risk of 24-week confirmed disability on EDSS. Progression by this definition was reduced in the DAC HYP 150 mg group (p = 0.0037) but not in the DAC HYP 300 mg group (p = 0.1487), compared with placebo. Hazard ratios for 24-week confirmed progression, relative to placebo, were 0.24 (95% CI: 0.09 to 0.63) for DAC HYP 150 mg and 0.60 (95% CI: 0.30 to 1.20) for DAC HYP 300 mg. The results in the 150 mg group cannot be considered significant because there has been no correction for multiplicity and, by the closed testing procedure used for major endpoints, results with 150 mg were only to be considered valid if the 300 mg dose group showed a significant effect.

Subgroup analyses

The sponsor performed a number of subgroup analyses of the primary endpoint (ARR), as shown in the Forest plot below (see Figure 6). A broadly consistent benefit for ARR was observed in most subgroups and despite the reduction in statistical power that comes from analysing subgroups, most comparisons with placebo remained significant. The exceptions were older subjects (age > 35 years) and subjects with previous disease-modifying treatment, where the 95% CIs for the rate ratios extended above unity indicating a non-significant result; however, even for these subgroups, the trends were in favour of active treatment. Reassuringly, a significant benefit was observed in subjects with and without relapses in the previous 12 months, with and without Gd enhancing lesions, and in subjects with both low and high EDSS (≤ 2.5 or > 2.5).

Figure 6. Annualised relapse rate by subgroup, Study 205MS201



NOTE: Rate ratios and 95% CI based on negative binomial regression model adjusted for the number of relapses in the past year except for the number of relapses in the past 12 months. In this model no adjustment was made. Only relapses confirmed by the INEC are included in the analysis. (a) Rate ratios and 95% CI based on poisson regression model adjusted for the number of relapses in the past year. (b) Rate ratios and 95% CI based on negative binomial regression model adjusted for the number of relapses in the past year. (c) a negative binomial regression model was used but no adjustment was made. (d) Subjects are defined as having high disease activity if they have >=2 relapses in the year prior to randomization and >=1 baseline Gd lesion.

Subgroup analysis for high disease activity versus low disease activity

A subgroup analysis based on high disease activity versus low disease activity was the subject of one of the supplementary data submissions. Even prior to receiving this suggestion for

additional analyses from the EMA, the sponsor had already conducted and submitted their own post-hoc analysis for high disease activity versus low disease activity subgroups.

In the study report for Study 205MS201, this analysis was summarised as follows:

'As a post-hoc analysis, the efficacy of DAC HYP was also evaluated in subjects with high disease activity at baseline, defined as ≥ 2 relapses in (the) year prior to randomisation and ≥ 1 Gd-enhancing lesion at baseline as well in subjects with and without prior MS treatment experience (excluding steroids). For the analysis in subjects with and without high disease activity, subgroups were evaluated for the primary endpoint (ARR), key secondary endpoints (number of new or enlarging T2 lesions), and tertiary endpoints (number of Gd lesions at Week 52, disease progression) and the DAC HYP 150 and DAC HYP 300 mg groups were combined. The percentage reduction in annualized relapse rate among those with high-disease activity was 51% (95% CI: 5.5% to 74.1%) compared to 51% (95% CI: 31.7% to 65.5%) in the low-disease activity group (see Figure 6 above). Across the other endpoints, DAC HYP demonstrated similarly high efficacy in subjects with both high- and low-disease activity prior to study entry.'

This analysis suggested that, in terms of the proportional reduction in relapse rate, DAC HYP has similar relative efficacy in subjects with both high and low disease activity (reducing ARR by 51% in both subgroups). The absolute benefit, in terms of number of relapses prevented, is expected to be higher in subjects with high disease activity.

Secondary and tertiary endpoints were also assessed according to baseline disease activity, as shown in the tables below. For disease progression, a numerical benefit was observed in both high-activity and low-activity subgroups, but statistical significance was only demonstrated for the low-activity subgroup with the DAC HYP dose groups pooled (hazard ratio = 0.54, 95% CI: 0.30 to 0.97; p = 0.0383). The analysis in the high-activity subgroup was underpowered, with only 30 placebo recipients in the ITT population and 58 subjects in the combined active groups. Active treatment was associated with a superior hazard ratio, and only one high-activity patient progressed on active treatment, compared to four on placebo, so the failure to achieve statistical significance could reflect low patient numbers (hazard ratio = 0.12, 95% CI: 0.01 to 1.07; p = 0.0574).

For radiological endpoints, the results were strong in both high-disease-activity and low-disease-activity subgroups. For both new T2 lesions and Gd lesions, the benefit with active treatment was consistent across both subgroups, and remained statistically significant for dose-pooled DAC HYP data ('DAC Total') in both subgroups, as shown in Tables 8 to 10 below.

Table 8. Time to sustained progression by baseline disease activity and treatment, 205MS201

	High Baseline Disease Activity		Low Baseline Disease Activity	
	Placebo	DAC Total	Placebo	DAC Total
Number of subjects in ITT population	30 (100)	58 (100)	165 (100)	341 (100)
Number of subjects who progressed	4 (13)	1 (2)	21 (13)	24 (7)
Time (wk) to progression (a)				
25th percentile	NA	NA	NA	NA
50th percentile	NA	NA	NA	NA
Estimated proportion of subjects with progression at 52 weeks (a)	0.138	0.018	0.133	0.076
Hazard Ratio (b)		0.12		0.54
(95% CI)		(0.01-1.07)		(0.30-0.97)
p-value vs placebo (b)		0.0574		0.0383
P-value for interaction (c)				0.2163

NOTE: Sustained progression of disability is defined as at least a 1.0 point increase on the EDSS from a baseline EDSS >=1.0 sustained for 12 weeks or at least a 1.5 point increase on the EDSS from a baseline EDSS of 0 sustained for 12 weeks.
(a) Estimated time to progression and proportion of subjects with progression based on the Kaplan-Meier product limit method.
(b) Estimated from a Cox proportional hazards model in each disease activity group (high or low) with covariates for baseline EDSS (<=2.5 versus >2.5) and treatment group.
(c) Estimated from a Cox proportional hazards model with covariates for baseline EDSS (<=2.5 versus >2.5), treatment group and the baseline EDSS by treatment group interaction.

Table 9. Number of new T2 lesions by baseline disease activity and treatment, Study 205MS201

	High Baseline Disease Activity		Low Baseline Disease Activity	
	Placebo	DAC Total	Placebo	DAC Total
Number of subjects in ITT population with non missing data	30 (100)	58 (100)	165 (100)	341 (100)
Number of new or newly enlarging T2 hyperintense lesions at 52 weeks				
0	1 (3)	15 (26)	37 (22)	180 (53)
1	0	7 (12)	16 (10)	42 (12)
2	2 (7)	9 (16)	12 (7)	41 (12)
3	2 (7)	10 (17)	7 (4)	25 (7)
>=4	25 (83)	17 (29)	93 (56)	53 (16)
n	30	58	165	341
Mean	14.1	4.4	7.2	2.5
SD	12.35	6.52	8.29	6.87
Median	12.0	2.1	6.0	0.0
Min, Max	0, 56	0, 33	0, 43	0, 86
Adjusted mean (a)	15.79	3.81	6.71	1.84
(95% CI)	(10.70, 23.29)	(2.83, 5.14)	(5.34, 8.44)	(1.54, 2.19)
Percent reduction (a)		75.86		72.63
(95% CI)		(59.97, 85.44)		(63.41, 79.53)
P-value versus placebo (a)		<.0001		<.0001
P-value for interaction (b)				0.1794

Note : Numbers in parentheses are percentages.
(a) Estimated from a negative binomial regression within each subgroup adjusted for baseline T2 lesions.
(b) Interaction p-value was estimated from a negative binomial regression model with covariates for baseline T2 lesions, treatment, baseline disease activity (high v low) and the disease activity by treatment interaction.

Table 10. Number of Gd-enhancing lesions by disease activity and treatment, Study 205MS201

	High Baseline Disease Activity		Low Baseline Disease Activity	
	Placebo	DAC Total	Placebo	DAC Total
Number of subjects in ITT population	30 (100)	58 (100)	165 (100)	341 (100)
Number of new Gd-enhancing lesions				
0	10 (33)	45 (78)	87 (53)	302 (89)
1	5 (17)	7 (12)	42 (25)	26 (8)
2	4 (13)	5 (9)	11 (7)	6 (2)
>=3	11 (37)	1 (2)	25 (15)	7 (2)
n	30	58	165	341
Mean	3.0	0.4	1.2	0.2
SD	3.72	0.96	1.80	0.75
95% CI for mean	1.62, 4.40	0.18, 0.69	0.88, 1.43	0.16, 0.32
Median	1.7	0.0	0.0	0.0
Min, Max	0, 14	0, 6	0, 10	0, 7
Odds ratio (a)		0.11		0.14
(95% CI)		(0.04, 0.28)		(0.09, 0.22)
p-value		<.0001		<.0001
P-value for interaction (b)				0.4621

Note : Numbers in parentheses are percentages.
(a) Estimated from an ordinal logistic regression model within each subgroup adjusted for baseline Gd lesions.
(b) Interaction P-value was estimated from an ordinal logistic regression with covariates for baseline Gd lesions, treatment, baseline disease activity (high v low), and the disease activity by treatment interaction

7.1.1.13. Overall conclusions for Study 205MS201

Overall, this placebo-controlled study was adequately designed and it used entry criteria and endpoints typical of other studies seeking to register disease-modifying agents in MS. The reduction in ARR was 50 to 54% across the two active dose groups, relative to placebo, without any apparent dose trend. A broadly consistent benefit was observed in a number of subgroups based on gender, age, EDSS status and measures of disease activity including a post-hoc definition of disease activity. A favourable trend was observed on disease progression, with nominally significant p-value in one dose group (the proposed dose group, but not the primary dose group in the statistical analysis plan).

The main points of contention between the current evaluator and the sponsor are as follows:

- The evaluator does not believe this study justifies the proposed indication because the entry criteria explicitly restricted the study to subjects with RRMS, whereas the indication refers to 'relapsing forms of MS'.
- The evaluator does not accept the sponsor's calculation of the relative risk reductions for the proportions of subjects that relapsed and the proportions of subjects that progressed, because these were based on instantaneous hazard ratios rather than on the actual proportions that relapsed and progressed over the period of study.
- The evaluator notes that, according to the prospective statistical analysis plan, this study did not show a significant effect on disease progression or MSIS-29 scores, whereas the proposed PI appears to indicate that the results for these endpoints were significant.

7.1.2. Study 205MS301

Study abstract: A Multicenter, Double-blind, Randomized, Parallel-group, Monotherapy, Active-control Study to Determine the Efficacy and Safety of Daclizumab High Yield Process (DAC HYP) versus Avonex (Interferon β-1a) in Patients with Relapsing-Remitting Multiple Sclerosis.

7.1.2.1. Study design, objectives, locations and dates

This Phase 3 study compared DAC HYP at the proposed dose (150 mg SC every 4 weeks) with an active control, IFN β-1a (tradename: Avonex) intramuscularly (IM) weekly, in subjects with

RRMS, using a randomised, double-blind, parallel-group design. The study was reasonably large (n = 1841) and it had an acceptable duration (up to 144 weeks), so it can therefore be considered the main pivotal study of the submission. The study lacked a placebo group, but the Phase 2 study 205MS201 provided placebo-controlled data and, in combination, the two studies provide a reasonably clear assessment of the efficacy of DAC HYP.

The primary objective of the study was *'to test the superiority of DAC HYP compared to IFN β-1a in preventing MS relapse in subjects with RRMS.'*

Secondary objectives were *'to test the superiority of DAC HYP compared with IFN β-1a in slowing functional decline and disability progression and maintaining quality of life in this subject population.'*

Additional objectives were to assess other long-term efficacy measures including neurological function and brain atrophy, to assess safety and tolerability, to gather pharmacokinetic (PK) data, and to study the effect of DAC HYP on pharmacodynamic (PD) markers.

The study was conducted at 246 sites in 28 countries (Argentina, Australia, Brazil, Canada, Czech Republic, Denmark, Finland, France, Georgia, Germany, Greece, Hungary, India, Ireland, Israel, Italy, Mexico, Moldova, Poland, Romania, Russia, Serbia, Slovenia, Spain, Sweden, Switzerland, Ukraine, United Kingdom, the United States of America) and ran from 11th May 2010 (first treatment) to 5th March 2014.

The most important inclusion and exclusion criteria, as listed in the study synopsis, are reproduced below. The complete list of formal entry criteria included more detailed restrictions based on concomitant diseases and abnormal laboratory tests at baseline

The inclusion criteria were similar to the other pivotal study, 205MS201, and attempted to restrict the study to subjects with active RRMS and no major progression.

The indicators of active disease were similar in both studies, but Study 205MS301 (discussed here) required 2 relapses in the last 3 years (or radiological evidence of multiple relapses) as well as 1 relapse in the last 12 months. EDSS restrictions were identical to the previous study (0.0 to 5.0 inclusive), and pose the same difficulty of interpretation in that subjects with higher EDSS may have had some progression between relapses, and therefore may have had SPMS or relapsing progressive MS. As noted in the discussion of Study 205MS201, SPMS and RPMS were notionally listed as exclusion criteria, but their definitions involved 3 months of continuous worsening, which may be very difficult to identify in clinical practice.

The requirement for at least 2 relapses in the previous 3 years (or a radiological substitute for clinical relapses) means that, despite the potential inclusion of some subjects with an element of progression, the cohort studied clearly had active, relapsing disease. The protocol-specified exclusion of subjects with SPMS and PRMS means that the study was focussed on the frequently relapsing, minimally-progressive end of the MS spectrum. It is therefore not known whether the benefits observed in this study would be reproduced in subjects with only one relapse in the previous 3 years, or in subjects where definite progression was present. This means that the proposed indication in the PI is too broad, as already discussed in the context of Study MS201 (see Section: Study 205MS201; Study design, objectives, locations and dates). This issue is addressed further in the discussion of the sponsor's response to key EMA Questions and in suggested edits to the PI.

Inclusion criteria

As per study report, the main inclusion criteria were:

- *'Ability to understand the purpose and risks of the study and provide signed and dated informed consent and authorisation to use protected health information in accordance with national and local subject privacy regulations.*
- *Must have been 18 to 55 years of age, inclusive, at the time of consent.*

- *Must have had a confirmed diagnosis of RRMS, as defined by McDonald criteria 1 through 4.*
- *Must have had an EDSS score between 0.0 and 5.0, inclusive.*
- *Must have experienced 2 or more clinical relapses within the previous 3 years, with at least 1 clinical relapse having occurred within the 12 months prior to randomisation or 1 or more clinical relapses and 1 or more new MRI lesions (Gd-enhancing and/or T2 hyperintense lesion) within the previous 2 years and with at least 1 of these events in the 12 months prior to randomisation. The new MRI lesion must have been distinct from one associated with the clinical relapse. The baseline MRI could be used to satisfy this criterion.*
- *Women of childbearing potential must have been willing to practice effective contraception during the study and been willing and able to continue contraception for 4 months after their last dose of study treatment.'*

Exclusion criteria

The main exclusion criteria were:

- *'Diagnosis of primary progressive, secondary progressive or progressive relapsing MS. These conditions required the presence of continuous clinical disease worsening over at least 3 months. Subjects with these conditions may also have had superimposed relapses, but were distinguished from subjects with RRMS by the lack of clinically stable periods or clinical improvement.*
- *Known intolerance, contraindication to, or history of noncompliance with IFN-β (Avonex) 30 µg. (Note: Current or prior use of an approved IFN-β preparation for MS, including Avonex, was allowed as long as the subject was appropriate for IFN-β treatment according to local prescribing information).*
- *History of abnormal laboratory results that, in the opinion of the investigator, were indicative of any significant cardiac, endocrine, haematological, hepatic, immunologic, metabolic, urologic, pulmonary, gastrointestinal, dermatologic, psychiatric, renal, neurological (other than MS), and/or other major disease that would have precluded administration of DAC HYP or IFN-β.*
- *An MS relapse that had occurred within the 50 days prior to randomisation and/or the subject had not stabilised from a previous relapse prior to randomisation.*
- *Any previous treatment with daclizumab or other anti-CD25 monoclonal antibody.*
- *Prior treatment with mitoxantrone, cyclophosphamide, fingolimod, or natalizumab within 1 year prior to randomization.'*

7.1.2.2. Study treatments

Subjects were randomised in a 1:1 ratio to (DAC HYP) or IFN β-1a (Avonex).

- DAC HYP was administered at the proposed dose of 150 mg SC 4-weekly.
- IFN β-1a was supplied as Avonex in a pre-filled syringe (PFS) for IM injection. Each 0.5 mL of comparator study drug contained 30 µg of IFN β-1a, and was administered once weekly. This is the standard registered dose for Avonex.

Subjects received study treatment in this study for up to 144 weeks. The study duration was described by the sponsor as '3 years' but this would require a treatment period of 156 weeks. By design, as the study was permitted to end when the last enrolled subject had received treatment for 96 weeks, some subjects enrolled later in the study received treatment for less than 2 years, even without considering those who terminated prematurely. The PI describes the study duration as 'a minimum of 2 to a maximum of 3 years (96 to 144 weeks).' This is not accurate, and the statement should be corrected.

It should be noted that, although Avonex is a registered active treatment for MS, it is considered by many neurologists to be less effective than other active treatments, and thus represents a soft target for head-to-head trials. Some direct head-to-head trials have suggested superiority of other interferon therapies relative to Avonex.[6, 7] It therefore appears plausible that DAC HYP might have shown less relative benefit if compared with a different active therapy.

The use of concomitant medications was restricted, as described previously for Study MS201: methylprednisolone was permitted for relapses, symptomatic treatments were stabilised prior to the randomised study period, where possible, and disease-modifying agents were prohibited.

7.1.2.3. Efficacy variables and outcomes

The study synopsis listed the following efficacy variables:

- *'Clinical outcomes:*
  - *Clinical relapses.*
  - *EDSS*
  - *Subject global assessment (as measured by the QoL questionnaires: EQ-5D and MSIS-29)*
  - *Multiple Sclerosis Functional Composite (MSFC)*
  - *Visual Function Test (VFT)*
  - *Symbol Digit Modalities Test (SDMT)*

    Relapses that were determined to meet protocol-defined criteria were subsequently evaluated by the (INEC)
- *Brain MRI outcomes:*
  - *Total number of new Gd-enhancing lesions*
  - *New or newly enlarging T2 hyperintense lesions*
  - *New T1 hypointense lesions*
  - *Volume of T2 hyperintense lesions*
  - *Volume of T1 hypointense lesions*
  - *Brain atrophy.'*

Primary efficacy endpoint

The primary efficacy endpoint was the adjusted ARR based on INEC-confirmed relapses. This is the same primary endpoint as the other pivotal DAC HYP study, and similar relapse-based primary endpoints have been used for most registration studies for disease-modifying agents in MS.

Secondary efficacy endpoints

Secondary efficacy endpoints were listed as follows, in rank order:

- Number of new or newly enlarging T2 hyperintense lesions on brain MRI over 96 weeks

[6] Durelli L et al; Independent Comparison of Interferon (INCOMIN) Trial study group. Every-other-day interferon beta-1b versus once-weekly interferon beta-1a for multiple sclerosis: results of a 2-year prospective randomised multicentre study (INCOMIN). Lancet. 2002 Apr 27;359(9316):1453-60.
[7] Schwid S et al. Full results of the Evidence of Interferon Dose-Response-European North American Comparative Efficacy (EVIDENCE) study: a multicenter, randomized, assessor-blinded comparison of low-dose weekly versus high-dose, high-frequency interferon beta-1a for relapsing multiple sclerosis. Clin Ther. 2007; 29(9):2031-2048.

- Proportion of subjects with confirmed disability progression defined by at least a 1.0-point increase on the EDSS from a baseline EDSS ≥ 1.0 that was sustained for 12 weeks or at least a 1.5-point increase on the EDSS from a baseline EDSS = 0 that was sustained for 12 weeks
- Proportion of subjects who were relapse free
- Proportion of subjects with a ≥7.5-point worsening from baseline in the MSIS-29 Physical Impact score at 96 weeks

These endpoints are reasonable. The EDSS is a standard measure of disability in MS, and it has been widely used and validated. It was used in this study as a baseline stratification measure, and it was used to define disease progression. The EQ-5D is a validated QoL measure used in many major efficacy studies.

The MSIS-29 is a validated, MS-specific QoL measure that was also used in Study 205MS201. It includes 2 sub-scales: the 20-item Physical Impact scale and the 9-item Psychological Impact scale. Increased scores represent worsening from baseline and decreased scores represent improvement. In validation studies, a change of ≥ 7.5 points was considered to be clinically meaningful.[8,9] This study used a ≥ 7.5 point change on the Physical subscale as a secondary endpoint.

The MSFC is a validated measure of disability that can be used as an alternative to the EDSS; it is based on aggregate performance in a number of tasks assessing walking, upper limb function and vision. MSFC scores are scaled by the standard deviation of raw scores obtained in a control group, so the resulting scores are somewhat abstract, difficult to read at a glance, and cannot be applied in isolation to a single patient. These are some of the reasons that the MSFC has not replaced the EDSS, which remains the gold standard measure of disability in MS.

The VFT and SDMT are validated, task-specific tests looking at a subset of neurological skills; they are not relevant unless subjects develop deficits in the domains under consideration, so they are only useful as minor endpoints.

The radiological outcomes listed above are standard objective measures of disease activity, with Gd lesions representing reasonably specific evidence of recent inflammatory activity, and hence active plaques. T2 lesions are the hallmark of MS, but require comparison with old scans to determine whether they are new or recent. T1 hypointense lesions represent focal loss of brain tissue, particularly axons, and correlate with permanent disability. Brain atrophy worsens throughout the course of the disease, and corresponds with long-term disease activity and cognitive decline; one issue posed by interpreting atrophy is that a reduction in inflammation in the brain may cause a reduction in volume, known as pseudoatrophy, and this may mask relative changes in the progression of true atrophy.

Although MRI outcomes are objective, they are usually treated as secondary endpoints because it is possible that a treatment might improve MRI measures without an associated clinical correlate, and such a treatment would not be clinically useful.

7.1.2.4. Randomisation and blinding methods

Randomisation was performed with an IVRS, and randomisation codes were not made available to patients or treating or rating clinicians. Randomisation to each of the two treatment arms was performed with a 1:1 ratio.

[8] Costelloe L, O'Rourke K, Kearney H, et al. The patient knows best: significant change in the physical component of the Multiple Sclerosis Impact Scale (MSIS-29 physical). J Neurol Neurosurg Psychiatry. 2007;78(8):841-4.
[9] Phillips GA, Wyrwich KW, Guo S, et al. Responder definition of the Multiple Sclerosis Impact Scale physical impact subscale for patients with physical worsening. Mult Scler. 2014.

Blinding was attempted by using a double-dummy approach, with a placebo for IFN-β-1a and a placebo for DAC HYP, each administered with the same dosing regimen as the corresponding active treatment. There was appropriate separation of the treating and rating neurologists, and the reporting radiologists, as described for Study 205MS201.

IFN-β-1a, like other interferon beta therapies, is associated with a number of 'tell-tale' side effects including injection site reactions (ISRs) and flu-like symptoms, also characterised as influenza-like illness. The IM approach means that ISRs are usually much less evident with Avonex than with some other IFN therapies, which are administered subcutaneously. Skin reactions were more commonly seen in the DAC HYP group, as acknowledged by the sponsor:

'In Study 205MS301, the incidence of cutaneous events by system organ class (SOC) was higher in the DAC HYP group (37%) than in the IFN β-1a group (19%)'.

These percentages do not appear to have included injection site reactions:

'The most common (cutaneous) events (≥ 2%) by preferred term in the DAC HYP group were rash (7%); eczema (4%); seborrheic dermatitis, acne, erythema, and pruritus (3% each); and dermatitis, dermatitis allergic, dermatitis contact, dermatitis atopic, rash maculopapular, dry skin, alopecia, urticaria, and psoriasis (2% each).'

Injection site-related AEs were nonetheless common:

'The overall incidence of AEs at the injection site was similar between the 2 groups (18% IFN β-1a versus 17% DAC HYP), as were the most common injection site AEs: injection site pain (11% versus 10%), injection site erythema (5% versus 4%), and injection site bruising (3% versus 2%).'

The incidence of ISRs appeared broadly balanced across the two groups, but the fact that the two treatments used different dosing sites and regimens means that any reaction at the site of an active injection is likely to have led to unblinding of the patient. Injections sites were covered during assessments by the rating neurologist (responsible for EDSS and relapse assessments), so this is not expected to have had a substantial effect on major efficacy endpoints.

The potential for unblinding due to FLS/ILI was anticipated by the sponsor, and some attempt was made to minimise this problem. Subjects received prophylactic treatment for FLSs, described in the study report as follows:

'In order to relieve flu-like symptoms for the first 24 weeks of study treatment dosing, all subjects were instructed to take acetaminophen (paracetamol) or ibuprofen or other nonsteroidal, anti-inflammatory drugs (NSAIDs) such as naproxen or aspirin prior to each Avonex (or matching placebo) injection and for the 24 hours after each injection at the recommended dose and frequency per the local labels. Additional doses of these protocol-designated products could be taken after 24 hours post-injection within the maximum daily dose recommended per local labels. After 24 weeks, the products could be discontinued at the discretion of the investigator.'

It is unlikely that these measures were sufficient to prevent some unblinding. In usual neurological practice, many subjects receiving IFN report flu like symptoms despite the use of prophylactic medications, and some subjects have ongoing flu like symptoms for more than 24 hours after each injection. It seems likely that, even if the prophylactic treatment was 100% effective in masking flu-like symptoms, subjects would occasionally forget doses (or omit them on purpose to determine if they were necessary). Furthermore, the protocol allowed subjects to cease prophylactic agents at 24 weeks, at which time it is possible that subjects who were receiving IFN would be exposed to flu-like symptoms for the first time even if they had enjoyed 100% mitigation of FLSs with prophylactic agents prior to that. The fact that the protocol allowed prophylactic agents to be dropped at 24 weeks 'at the discretion of the investigator' highlights the fact that flu-like symptoms and the adequacy and necessity for prophylaxis were explicitly discussed by the patient and the treating neurologist, leading to potential unblinding of the neurologist as well as the patient.

An assessment of the incidence of influenza like illness, reported as an AE shows that it was seen in a substantial portion of IFN β-1a recipients (38%), and was much less commonly observed in DAC HYP recipients (10%). According to the sponsor:

'Across the study period, 346 subjects in the IFN β-1a group and 88 subjects in the DAC HYP group reported at least 1 event of influenza-like illness.'

The risk that this tell-tale side effect could have led to unblinding would have been increased by the fact that the two drugs had different dosing schedules. A patient who experienced flu-like symptoms after every weekly IM injection could easily guess they were receiving IFN β-1a, and subjects who experienced similar symptoms every 4 weeks after a SC injection could guess they were receiving DAC HYP.

The evaluator found no evidence that the sponsor took steps to assess the extent of unblinding (this could have been achieved by asking subjects and neurologists to guess the treatment assignment at the end of the study).

A digital search of the case study report for 'unblinding' reveals that there were 6 instances of accidental blinding due to logistical errors in which treatment assignments were revealed, but there is no mention of any attempt to quantify the extent of unblinding due to side effects. This is a substantial methodological flaw in the study; on the other hand, the sponsor performed sensitivity analyses in subjects with and without flu like symptoms, and this analysis indirectly suggests that inadvertent unblinding, although likely to be present in a substantial number of patients, did not play a large role in determining the outcome.

7.1.2.5. Analysis populations

The sponsor defined three analysis populations.

- The ITT population included all randomised subjects who received at least 1 dose of any study treatment, analysed according to the group to which they were randomised.
- The per-protocol population was defined as all subjects from the ITT population who satisfied the following conditions:
    a. Met both inclusion criteria related to MS-specific disease activity:
        i. Had a confirmed diagnosis of RRMS according to McDonald criteria 1-4 and a cranial MRI demonstrating lesion(s) consistent with MS.
        ii. Had a baseline EDSS between 0.0 and 5.0, inclusive.
    b. Compliant with study treatment: ≥ 90% of DAC HYP or Avonex doses up to Week 96.
    c. Did not permanently discontinue study treatment prior to Week 96.
- The safety population comprised all subjects who received at least 1 dose of study any treatment.

For the primary efficacy analysis and most secondary endpoints, the main analysis was performed on the ITT Population, and subjects were analysed in the group to which they were randomised.

For the number of new or newly enlarging T2 lesions on MRI, the analysis was based on the subset of subjects with a non-missing post-baseline assessment.

The per-protocol population was used for sensitivity analyses, and all major efficacy endpoints were reassessed in this population, generally producing results consistent with the ITT analysis.

For the safety analysis, subjects in the Safety Population were analysed according to the treatment they actually received.

Overall, the Sponsor's approach to these analysis populations was appropriate.

7.1.2.6. Statistical methods

Statistical methods in Study 205MS301 were similar to those in Study 205MS201, and they were broadly appropriate apart from a number of 'sensitivity analyses' that used questionable imputation methods to reanalyse the progression data, as well as a potentially misleading calculation of relative risk reduction for the proportion of subjects experiencing a relapse. These issues are discussed in more detail in the statistical methods and efficacy results of the previous study above.

The primary efficacy endpoint was the adjusted ARR based on INEC-confirmed relapses. The analysis of this endpoint included data from all ITT subjects until they completed the End of Treatment Period Visit, switched to alternative MS medication, or withdrew from the study. The difference in ARR between DAC HYP 150 mg and IFN β-1a was assessed with a negative binomial regression model, adjusting for baseline relapse rate, history of prior IFN β-1a use, baseline EDSS (≤ 2.5 versus > 2.5) and baseline age (≤ 35 versus > 35 years). This is very similar to Study 205MS201 but with one additional adjustment factor (history of IFN β-1a use).

Analysis methods for secondary endpoints included:

- negative binomial regression (for number of T2 hyperintense lesions)
- Cox proportional hazards and Kaplan-Meier product limit estimator (for disability progression as measured by an increase in EDSS score, and for proportions of subjects who were relapse free)
- logistic regression (for proportion of subjects with a ≥ 7.5-point worsening in the MSIS-29 Physical Impact score).

The main analyses of efficacy endpoints excluded data after subjects switched to alternative MS medications, but the sponsor performed additional sensitivity analyses that included data after switching.

To control for multiplicity of secondary endpoints, a closed testing procedure was used. Endpoints were ranked in terms of priority, and if statistical significance was not achieved for an endpoint, all endpoints of a lower rank were considered not statistically significant. The p-values presented were nominal results, not adjusted for multiplicity.

Tertiary endpoints did not include adjustments for multiple comparisons and endpoints.

7.1.2.7. Sample size

Sample size estimations were based on the primary endpoint ARR. Power was estimated from simulations that assumed a 21% drop-out rate, an average of 2.4 years of follow-up, and an ARR of 0.27 in the IFN β-1a group. With these assumptions, it was estimated that a sample size of 900 subjects per treatment group would have approximately 90% power to detect a 24% reduction in the ARR between the IFN β-1a treatment group and the DAC HYP treatment group, based on a negative binomial regression model with a standard 5% type 1 error rate (≤ 0.05). These calculations suggested that approximately 1800 subjects were required for the study and this target was exceeded (n = 1841).

Overall, these assumptions appear plausible, and the study showed itself to be adequately powered for this endpoint.

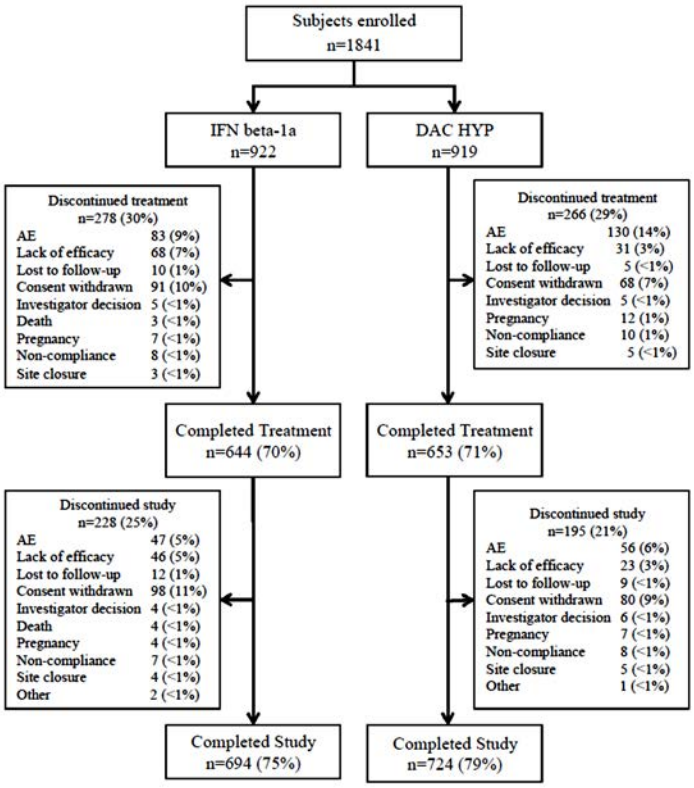
The study was not specifically powered for the key secondary endpoint of disability progression, and failed to show a significant benefit despite a slightly favourable trend.

7.1.2.8. Participant flow

Patient disposition is summarised in Figure 7 below. Completion rates were similar in both treatment groups (75% for IFN β-1a and 79% for DAC HYP) and were acceptable for a large, long study of this nature. The most common reasons for withdrawal were AEs, apparent lack of

efficacy, and withdrawal of consent. The reasons were broadly balanced across the two treatment groups, making it relatively unlikely that the study experienced major withdrawal bias. There was a slight excess of IFN β-1a subjects withdrawing due to a perceived lack of efficacy (7% compared to 3% in the DAC HYP group).

Figure 7. Subject disposition and participant flow, Study 205MS301



7.1.2.9. Major protocol violations/deviations

Major protocol deviations were summarised by the sponsor as follows:

'Overall, the incidence and category of major protocol deviations were similar between the two treatment groups. The most common major deviations (≥ 20%) were 'Informed Consent' (32% IFN β-1a versus 33% DAC HYP), 'Key Study Procedures' (28% IFN β-1a versus 27% DAC HYP), and 'Other' (27% (for) each group).'

This is suggestive of a high level of protocol deviations, though the description does not clearly indicate the nature of the deviations. The sponsor's text provided a link to a table of deviations, reproduced below, but this also lacked sufficient detail and it is not possible to determine whether the deviations related to 'Key Study Procedures' or 'Other' substantially compromised the study. The sponsor should be asked to clarify this issue.

Table 11. Summary of major protocol deviations, Study 205MS301

Summary of major protocol deviations

	IFN beta-1a 30 mcg	DAC HYP 150 mg
Number of subjects in the ITT population	922 (100)	919 (100)
Informed Consent	294 (32)	299 (33)
Eligibility	25 (3)	29 (3)
Study Tx Administration	144 (16)	148 (16)
Prohibited Con Med	44 (5)	52 (6)
Key Study Procedure	262 (28)	244 (27)
Other	253 (27)	244 (27)

NOTE: Numbers in parentheses are percentages.

7.1.2.10. Baseline data

The two treatment groups were well matched in terms of demographics. Although this was not demonstrated in a convenient table within the study report, the clinical overview included a one-page table covering key aspects of the demographics of both pivotal studies. (The relevant sections for Study 205MS301 are the last two columns of Table 12 below). Both treatment groups had a similar gender distribution, mean age, and racial mix (not shown in the table).

The treatment groups were also reasonably matched for baseline disease characteristics, including mean years since diagnosis (slightly longer in the DAC HYP group), relapses in the last 3 years (≤ 2 for 57% of patients in both groups, ≥ 3 for 47% of patients in both groups), relapses in the last 12 months, mean EDSS scores (close to 2.5 in both groups), and MRI lesion counts.

Overall, the balance between treatment groups was acceptable, and the results are unlikely to have been significantly influenced by unequal risks at baseline. Furthermore, the population studied was reasonably representative of the population likely to be considered for treatment with DAC HYP.

Table 12. Demographics and baseline disease characteristics (Study 205MS201/301)

Summary of the Demography and Baseline Disease Characteristics in Study 205MS201 and Study 205MS301

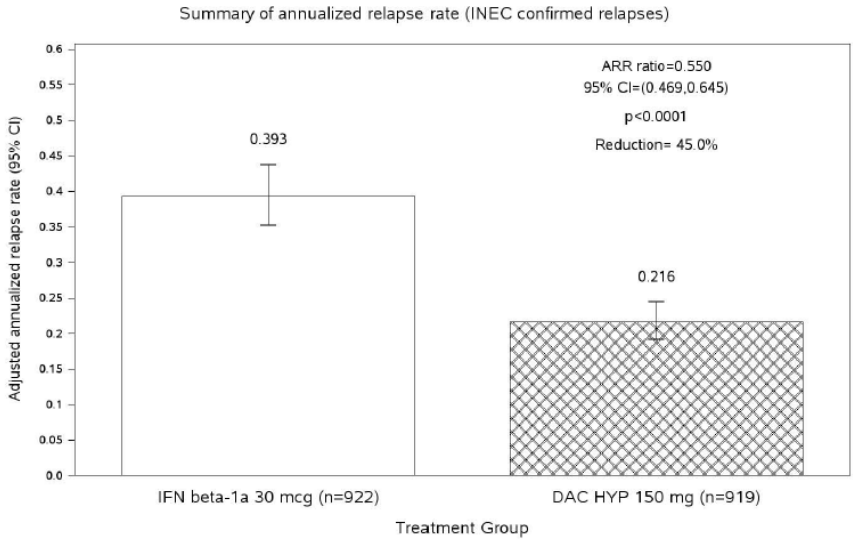
	Study 201		Study 301	
	Placebo	DAC HYP 150 mg	IFN beta-1a 30 mcg	DAC HYP 150 mg
Number of subjects randomized	204 (100)	208 (100)	922 (100)	919 (100)
Age (yrs)				
Mean (SD)	36.6 (9.02)	35.3 (8.94)	36.2 (9.32)	36.4 (9.36)
Female - no.(%)	128 (63)	140 (67)	627 (68)	625 (68)
Years since diagnosis				
Mean (SD)	4.1 (5.26)	4.5 (4.96)	4.1 (4.70)	4.2 (4.97)
No. Relapses in past 3 years - no.(%)				
n	204	208	922	918
<=2	132 (65)	123 (59)	524 (57)	524 (57)
>=3	72 (35)	85 (41)	398 (43)	394 (43)
EDSS Score				
Mean (SD)	2.73 (1.166)	2.81 (1.147)	2.54 (1.257)	2.48 (1.206)
Median	2.50	3.00	2.25	2.00
>=3.5 - no.(%)	75 (37)	85 (41)	291 (32)	260 (28)
Gd Lesions				
n	203	206	909	900
Mean (SD)	2.0 (4.48)	2.1 (3.47)	2.3 (5.85)	2.0 (5.86)
>=1 - no.(%)	90 (44)	106 (51)	414 (45)	398 (43)
T2 Lesions				
n	203	206	908	900
Mean (SD)	39.5 (32.17)	44.6 (34.71)	51.8 (37.39)	49.2 (35.52)
Prior ABCR or immunodulatory - no.(%)	38 (19)	50 (24)	376 (41)	380 (41)
Number of relapses in prior 12 months				
n	204	208	922	919
Mean (SD)	1.3 (0.60)	1.4 (0.73)	1.6 (0.75)	1.5 (0.72)

7.1.2.11. Results for the primary efficacy outcome

The study achieved a significant positive result for its primary endpoint with an adjusted ARR in the IFN β-1a group of 0.393 relapses/year, and a rate of 0.216 relapses/year in the DAC HYP group (95% CI: 0.353 to 0.438 in the IFN β-1a treatment group and 0.191 to 0.244 in the DAC HYP treatment group). Unadjusted ARRs were broadly similar (0.353 and 0.212 for IFN β-1a and DAC HYP, respectively).

These results correspond to a relative reduction of 45% in ARR ($p < 0.0001$) with DAC HYP, compared to IFN β-1a. The rate ratio of ARRs was 0.550 (95% CI: 0.469 to 0.645), indicating that a reduction in relapse rate of at least 35% could be expected with DAC HYP (based on the pessimistic upper limit of the 95% CI for the rate ratio) compared to an active treatment that has been shown to be superior to placebo. These are strong results for a head-to-head study and show a clear, clinically meaningful benefit with DAC HYP in reducing relapses. The expected benefit over placebo is not easily estimated from these figures, but would be expected to be better than the observed benefit over weekly IFN β-1a, and consistent with the results of Study 205MS201 (where a 54% reduction in ARR was observed for the 150 mg dose).

Figure 8. Annualised relapse rates, Study 205MS301

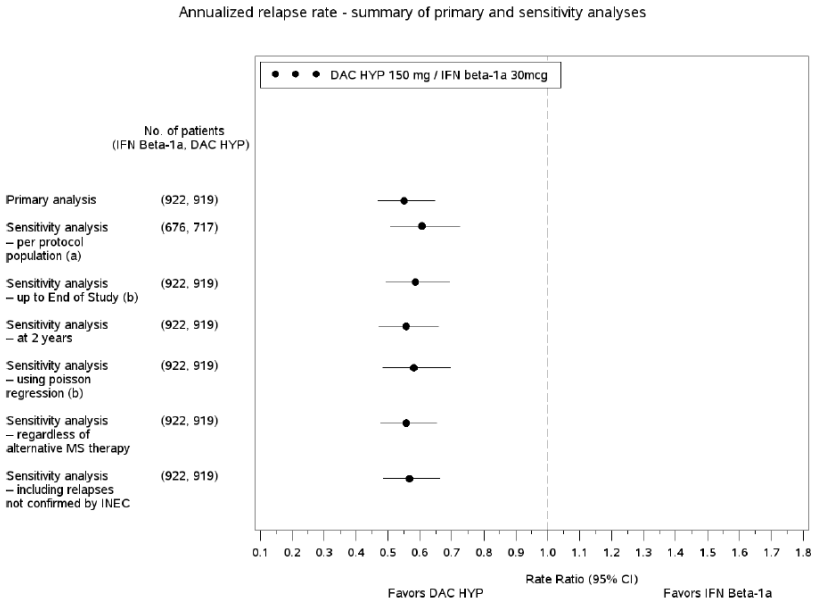


NOTE 1: Only relapses confirmed by INEC are included in the analysis.
2: Data after subjects switched to alternative MS medications are excluded.
3: Estimated from a negative binomial regression model adjusted for the baseline relapse rate, history of prior IFN beta use, baseline EDSS (<=2.5 vs >2.5) and baseline age (<=35 vs >35).
4: Annualized relapse rate (ARR) ratio with 95% CI and percent reduction are for DAC HYP 150 mg relative to IFN beta-1a 30mcg. P-value is from the likelihood ratio test that there is no difference between the two treatment groups.

Similar results were obtained in all pre-specified subgroups based on demographics and baseline disease characteristics (discussed under sub-analyses for this study).

Sensitivity analyses of this result, using different statistical approaches as shown in the figure below, produced broadly concordant results and suggest that the effect of DAC HYP on ARR was statistically and methodologically robust. In the per protocol population, the results were slightly inferior to the ITT results, but still consistent with a clear head-to-head benefit over IFN β-1a: DAC HYP reduced the ARR by 39% relative to IFN β-1a (rate ratio: 0.606 (95% CI: 0.508 to 0.724); $p < 0.0001$).

Figure 9. Annualised relapse rate: summary of primary and sensitivity analyses, Study 205MS301



(a) Estimated from a negative binomial regression model adjusted for the baseline relapse rate, history of prior IFN beta use, baseline EDSS (<=2.5 vs >2.5) and baseline age (<=35 vs >35).
(b) Estimated from a poisson regression model adjusted for the baseline relapse rate, history of prior IFN beta use, baseline EDSS (<=2.5 vs >2.5) and baseline age (<=35 vs >35). The model was adjusted for over-dispersion.

In a related analysis, restricted to severe or serious relapses, a broadly similar proportional benefit was observed, further supporting the conclusion that the reduction in relapses was clinically meaningful: DAC HYP produced a 38% reduction in severe or serious MS relapses compared with INF β-1a (p = 0.0021).

7.1.2.12. Results for other efficacy outcomes

The study specified four secondary efficacy endpoints, which were ranked as follows:

1. Number of new or newly enlarging T2 hyperintense lesions on brain MRI over 96 weeks
2. Progression of disability as measured by EDSS score
3. Proportion of subjects free from relapse
4. Proportion of subjects with a ≥ 7.5-point worsening from baseline in the MSIS-29 Physical Impact score at Week 96

T2 hyperintense lesions

The number of new or newly enlarging T2 hyperintense lesions at Week 96 was significantly and substantially reduced by DAC HYP, relative to IFN β-1a: the adjusted mean lesion count was 9.44 (95% CI: 8.46 to 10.54) in the IFN β-1a treatment group and 4.31 (95% CI: 3.85 to 4.81) in the DAC HYP treatment group. This amounts to a reduction of 54.4% (95% CI: 46.9% to 60.8%; p < 0.0001) with DAC HYP. Broadly similar results were obtained with a variety of sensitivity analyses (not shown in this evaluation report).

Progression of disability

Disability, as measured by the EDSS, was assessed at baseline and at all study visits throughout the treatment period. Progression was defined as a ≥ 1.0-point increase on the EDSS (or a ≥ 1.5 point increase on the EDSS from a baseline EDSS of 0) sustained for 12 weeks. The primary method of comparing treatment groups was based on a Cox proportional hazards model, adjusted for baseline EDSS (EDSS ≤ 2.5 versus EDSS > 2.5), history of prior IFN β use, and baseline age (age ≤ 35 versus age > 35 years).

By the primary prospectively specified analysis method, there was no significant difference between the groups: the hazard ratio for confirmed progression was 0.84 (DAC HYP/IFN β-1a), but the 95% CI included the possibility that progression was increased with DAC HYP (95% CI: 0.66 to 1.07).

The PI contains the following description:

'Zinbryta treated patients had a relative risk reduction in 12 week and 24 week confirmed disability progression of 16%, (95% CI: -7% to 34%; p = 0.16) and 27% (95% CI: 2% to 45%; p = 0.03) respectively compared to interferon beta-1a (IM) treated patients.'

This statement acknowledges that the results were not uniformly significant, but it fails to acknowledge that 12 week confirmed progression was a higher ranking endpoint than 24 week confirmed progression. The sponsor's study report was somewhat less clear, reporting that: 'DAC HYP reduced the risk of disability progression by 16% (p = 0.1575) compared with IFN β-1a'. It is important to note that the reported reduction of 16% is merely the central estimate of an uncertain range that included a 7% increase in progression.

As mentioned earlier in this report, relative risk estimates derived from hazard ratios do not necessarily reflect those derived from the actual proportions reaching a hazardous endpoint. For this particular endpoint, the distinction was numerically minor. The 16% reduction in risk cited by the sponsor has presumably been derived directly from the estimated hazard ratio of 0.84; it nonetheless appears to be consistent with a direct comparison of the overall adjusted progression rates at 96 weeks. As shown in Table 13 below, the adjusted proportions of progressed patients in each group at 144 weeks were 0.203 and 0.162 for the IFN β-1a and DAC HYP groups, respectively. This means that, at the 144-week time point, the DAC HYP progression rate was 79.8% (0.162/0.203) of the IFN β-1a rate. At the 96-week time point, the progression rate was 83.9% of the IFN β-1a rate (0.120/0.143), consistent with a 16% reduction.

Table 13. Time to 12-week sustained EDSS progression, Study 205MS301

Summary of time to 3-month sustained disability progression

	IFN beta-1a 30 mcg	DAC HYP 150 mg
Number of subjects in the ITT population	922 (100)	919 (100)
Number of subjects progressed	140 (15)	121 (13)
Time (weeks) to progression (a)		
10th percentile	60.1	72.6
25th percentile	NA	NA
50th percentile	NA	NA
Estimated proportion progressed (a)		
24 weeks	0.036	0.035
48 weeks	0.081	0.064
72 weeks	0.114	0.095
96 weeks	0.143	0.120
120 weeks	0.161	0.148
144 weeks	0.203	0.162
Hazard ratio (DAC HYP/ IFN beta-1a) and 95% CI (b)		0.84 (0.66, 1.07)
p-value vs IFN beta-1a (b)		0.1575

The sponsor performed a number of 'sensitivity analyses' of these results. The first of these to be described in the study report was as follows:

'In the primary analysis of 12-week confirmed disability progression, all subjects who had a tentative disability progression and did not have an available confirmatory assessment were assumed to be non-progressors and were censored at the time of the last assessment. A pre-specified sensitivity analysis of 12-week confirmed disability progression was performed based on the alternative assumption that confirmed disability progression would occur at a similar rate as

that for subjects who completed the confirmatory assessment in the trial (after adjustment for treatment group, baseline EDSS, change in EDSS at time of tentative progression, and presence of a relapse within the 29 days prior to the tentative progression). In this analysis, DAC HYP reduced the risk of 12-week confirmed disability progression by 21% as compared with the IFN β-1a group (hazard ratio (DAC HYP/IFN β-1a) of 0.79 (95% CI: 0.62 to 1.00; p = 0.0469))'.

This approach does not appear justified. By taking the rates of confirmed progression in each group and applying them (without supporting evidence) to additional cases of unconfirmed progression, the sponsor has allowed each case of confirmed progression to be counted as more than a single case: it is first counted as a confirmed case, and then its occurrence additionally affects the assumed rate of confirmation in different cases that did not actually reach confirmation. This double-accounting artificially inflates the statistical power of the analysis, and leads to a nominally significant p-value that does not reflect the true statistical uncertainty in the data.

Similar reasoning applies to another 'sensitivity analysis' described in the same paragraph of the clinical study report:

'An additional pre-specified sensitivity analysis was carried out in which all tentative progressions with no confirmation assessment were assumed to be confirmed… In this analysis, DAC HYP also significantly reduced the risk of 12-week confirmed progression by 24% compared with the IFN β-1a group (hazard ratio (DAC HYP/IFN β-1a) of 0.76 (95% CI: 0.61 to 0.95; p = 0.0157)).'

In this analysis, cases of progression would clearly be contaminated by the inappropriate inclusion of relapses, because the analysis method simply assumes that unconfirmed worsening in EDSS is always sustained. This circumvents the methodological processes originally designed to assess progression without having the assessment confounded by relapses.

The sponsor provided one additional analysis purporting to show that progression was significantly reduced:

'Additional related analyses also supported a significant treatment effect of DAC HYP in preventing disability progression compared with IFN β-1a. The risk of treatment failure (defined as the earliest of sustained progression of disability (at least a 1.0-point increase on the EDSS from a baseline EDSS ≥ 1.0 or at least a 1.5-point increase on the EDSS from a baseline EDSS of 0 that was sustained for 12 weeks), use of alternative MS medication, or treatment discontinuation due to lack of efficacy) was reduced by 19% in DAC HYP-treated subjects relative to IFN β-1a (hazard ratio (DAC HYP/IFN β-1a) of 0.81 (95% CI: 0.65 to 0.99; p = 0.0421)).'

This analysis refers to 'treatment failure' and the sponsor suggests that this is a potential surrogate for disease progression. In the sponsor's quoted paragraph, though, treatment failure is defined to include, not just cases of progression, but all subjects that switched MS therapies or withdrew due to lack of efficacy.

It seems inevitable that a high proportion of subjects switching therapy or discontinuing for lack of efficacy did so because of relapses, and so this composite endpoint conflates a treatment benefit for which there is already good evidence (reduced relapses on DAC HYP) with one for which there is no solid evidence (reduced disease progression on DAC HYP).

All of these 'sensitivity analyses' are rejected. It must be concluded that across the full study cohort disease progression was not significantly affected by treatment allocation. Note that the term 'sensitivity analysis' is usually used for situations where the result is so robust it survives reanalysis with pessimistic or conservative methodology, not situations where a negative result can be rendered positive through optimistic assumptions.

Contrary to these conclusions, the sponsor argues:

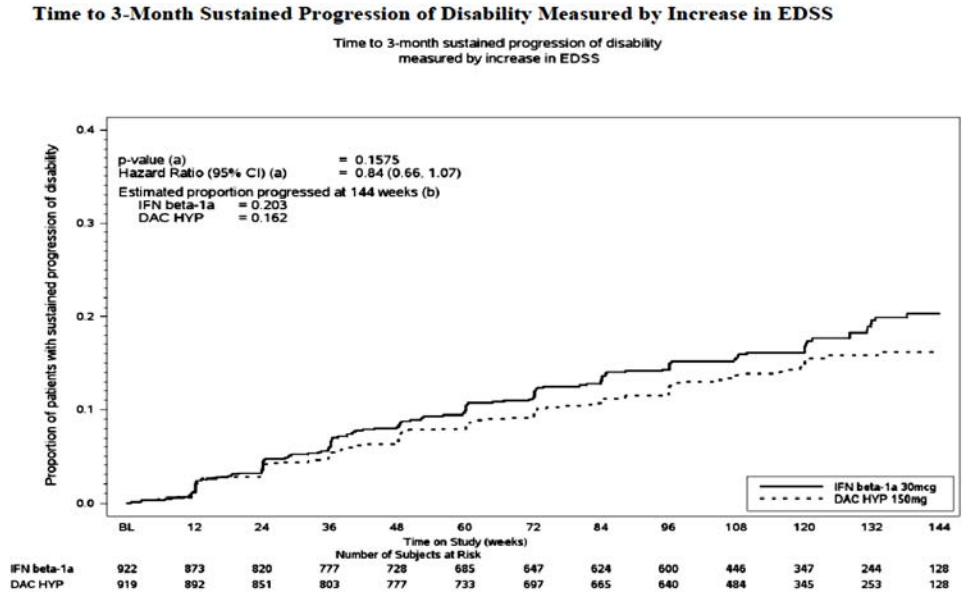
'While overall the primary and pre-specified sensitivity analyses were consistent with each other, the estimated treatment effects were stronger and reached statistical significance except when it

was assumed that disability progression did not occur in any subject who was censored after a tentative disability progression (the primary analysis). This assumption of the primary analysis did not appear to be valid because the risk of confirmed disability progression was substantial after a tentative disability progression among subjects with 3-month confirmatory visits (34% in the IFN β-1a group and 37% in the DAC HYP group). Censoring after a tentative disability progression was nearly twice as common in the IFN β-1a group as in the DAC HYP group (43 subjects versus 24 subjects, respectively), reflecting a proportionally higher number of tentative disability progressions in the IFN β-1a group. While the number of subjects censored after a tentative disability progression (n = 67) was small relative to the total number of subjects with a tentative disability progression in the trial (n = 736), the assumptions made about disability progression in these censored subjects affected whether the test of statistical significance for disability progression was above or below the 0.05 significance threshold. Given this imbalance between the treatment arms and the considerable risk for disability progression expected in these subjects, the primary analysis cannot be assumed to have provided an unbiased estimate of the treatment effect. Overall, based on the pattern of censoring and the high risk of confirmed disability progression after a tentative disability progression, the sensitivity analyses are considered most likely to have provided the most accurate estimate of the treatment effect on 12-week confirmed disability progression in this study.'

These comments indicate that only about one third of tentative progressions, if followed up, converted to confirmed progressions. They also indicate that censoring after a tentative progression was nearly twice as common in the IFN β-1a group (possibly in part because relapses were twice as common in that group), so this group is the one that would end up with more cases of 'assumed progression' by any of the imputation methods proposed. The Evaluator agrees that it cannot be assumed the primary analysis was unbiased, but it cannot be assumed that the suggested alternative analyses were unbiased, either – and in the case of imputing 100% conversion rates from 'tentative' to 'confirmed', the method clearly conflates relapses and progression. A bigger problem is that, even if the suggested imputation methods were unbiased, they artificially inflate the power of the analyses, and even then only just reach nominally significant p-values. If appropriate adjustments were made for the Sponsor's use of multiple analysis methods, it is likely that these additional analyses would no longer achieve even nominal significance.

In conclusion, this study did not demonstrate that DAC HYP prevents progression in comparison to IFN β-1a.

Figure 10. Time to sustained EDSS progression, Study 205MS301



When an additional analysis was performed using 24-week sustained progression (pre-specified as a tertiary endpoint), the difference between the groups emerged as statistically significant: relative to IFN β-1a, DAC HYP was associated with a reduced risk of 24 week confirmed disability progression, expressed as a hazard ratio of 0.73 (95% CI: 0.55 to 0.98; p = 0.0332). These results have not been adjusted for multiplicity and cannot be considered statistically robust.

Proportion of subjects free from relapse

The primary analysis of the secondary endpoint, was based on INEC-confirmed relapses in the ITT population occurring between the first dosing date and the subject's End of Treatment Visit or time of censoring. A total of 392 subjects (43%) in the IFN β-1a group and 260 subjects (28%) in the DAC HYP group had an INEC-confirmed relapse. The Kaplan-Meier estimate was used to assess the likely proportions of relapse-free subjects at different time points in the IFN β-1a and DAC HYP groups, respectively:

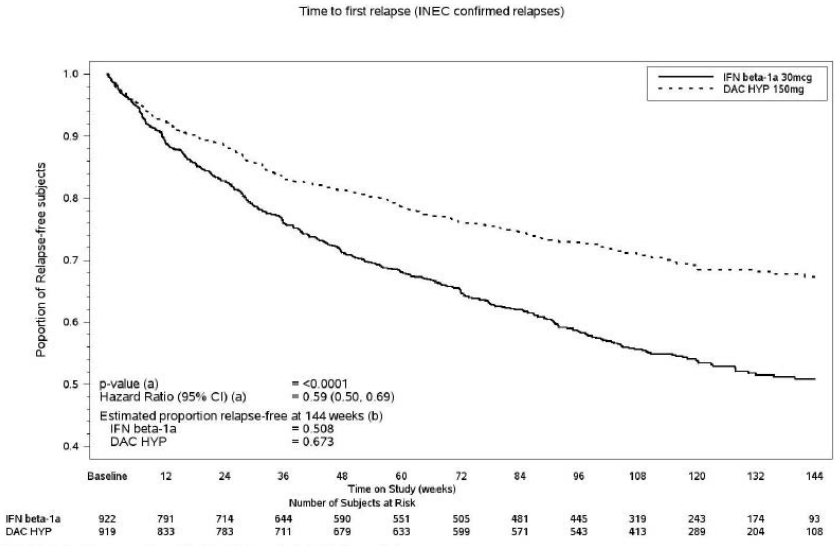
- 71.2% and 81.2% at 48 weeks;
- 58.5% and 72.9% at 96 weeks;
- 50.8% and 67.3% at 144 weeks.

The hazard ratio (DAC HYP/IFN β-1a) for the risk of relapse was 0.59 (95% CI: 0.50 to 0.69; p < 0.0001), which is a strong result for a head-to-head study, and consistent with the observed reduction in ARR.

Based on the proportions relapsing at 144 weeks (49.2% and 32.7% in the IFN β-1a and DAC HYP groups respectively), the relative proportion of relapsing subjects for DAC HYP was 66.5% (32.7/49.2) of the proportion observed with IFN β-1a. That is, the relative reduction in proportion relapsed was 33.5%. The sponsor's clinical study report claimed that these results demonstrated that the risk of relapse was reduced by 41% in the DAC HYP group, compared to IFN β-1a but this inflated estimate presumably refers to the reduction in instantaneous hazard (hazard ratio = 0.59), as already discussed in the context of study MS201. Patients and clinicians are more likely to be interested in the risk reduction over a defined time period, which can be estimated as 33.5% for a period of 144 weeks. A similar calculation suggests that the reduction in the proportion relapsing at 48 weeks was approximately 34.7%(The proportions relapsing were 28.8% and 18.8%, in the IFN β-1a and DAC HYP groups, respectively, and 18.8/28.8 is 65.3%). Overall, by these calculations, the reduction in the proportion of subjects relapsing was about 34-35% with DAC HYP, compared to IFN β-1a, rather than the 41% suggested in the clinical study report. The PI also includes the inflated estimate of 41%, and it should be changed to reflect the actual reduction in risk over the course of the study.

Similar results for reductions in proportions relapsing were obtained with a number of sensitivity analyses (not shown in this evaluation).

Figure 11. Time to first relapse, Study 205MS301



Proportion of subjects with a ≥ 7.5-point worsening from baseline in the MSIS-29 physical impact score at Week 96

The MSIS-29 Physical Impact score was assessed with a logistic regression model that included adjustments for the baseline Physical Impact score, baseline BDI, history of prior IFN β use, and baseline age (age ≤ 35 versus age > 35 years). Data imputation was required at Week 96 data for a substantial of patients: 202 subjects in the IFN β-1a group and 169 subjects in the DAC HYP group.

At the pre-specified main time point of 96 weeks, 213 subjects (23%) in the IFN β-1a group had a ≥ 7.5-point worsening from baseline, compared with 171 subjects (19%) in the DAC HYP group. The difference was statistically significant, with an odds ratio (DAC HYP/IFN-β 1a) of 0.76 (95% CI: 0.60 to 0.95; p = 0.0176).

A direct comparison of the proportions showing worsening (19% versus 23%) produces a ratio of 83% (19/23 = 0.826), or a relative improvement of approximately 17% for this endpoint. The absolute difference of 4% appears to be of rather modest clinical benefit.

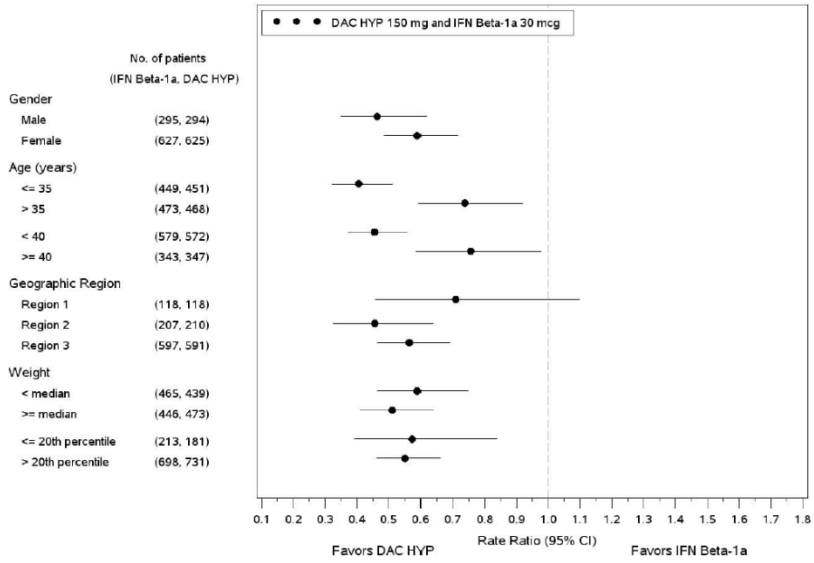
Sensitivity analyses of this endpoint, including assessments in the PP Population, produced broadly similar results.

Because IFN β-1a is associated a number of side effects including fatigue, flu-like malaise, spasm and depression, improvements in quality of life relative to interferon may partly reflect adverse effects of interferon rather than efficacy benefits of DAC HYP. The comparison with placebo, in the previous pivotal study, did not achieve significance.

Subgroup analyses

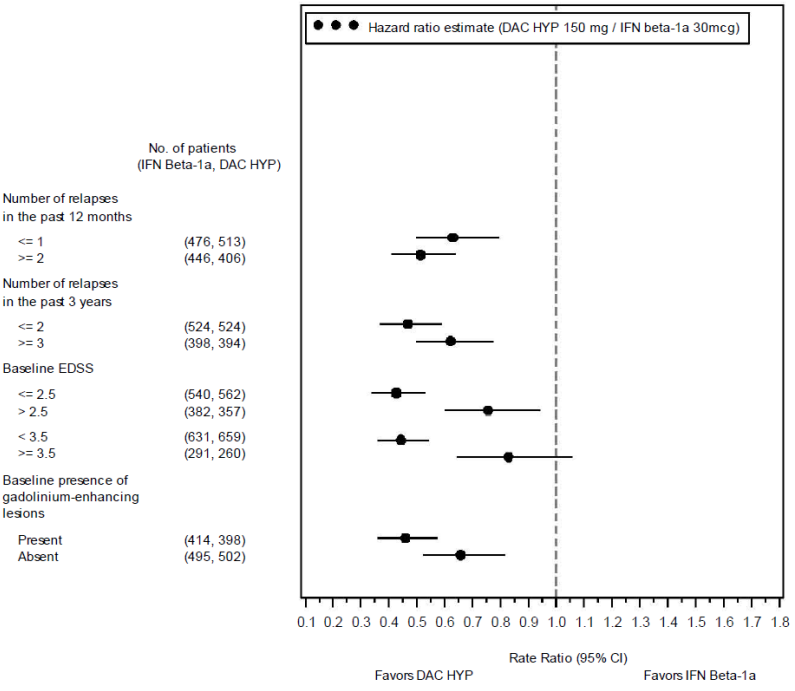
The sponsor performed subgroup analyses for the ARR, based on a number of different baseline prognostic factors including number of relapses in different time frames, EDSS scores, MRI characteristics, time since diagnosis and previous treatment. These are summarised in the Forest plot below (see Figure 12). In all but one subgroup, a significant treatment effect was observed, despite the fact that the subgroups had less subjects and the analysis had less power than the original analysis with the full cohort. The exception was the subgroup of patients with EDSS ≥ 3.5 at baseline. In this group, there was a trend to benefit with DAC HYP, but the hazard ratio estimate was less favourable than for other subgroups and the 95% CI included unity. This suggests that subjects with more advanced disease (who are more likely to have reached a SPMS stage of the illness rather than having pure RRMS) may be less responsive to treatment.

Figure 12. Annualised relapse rate by demographic subgroups, Study 205MS301



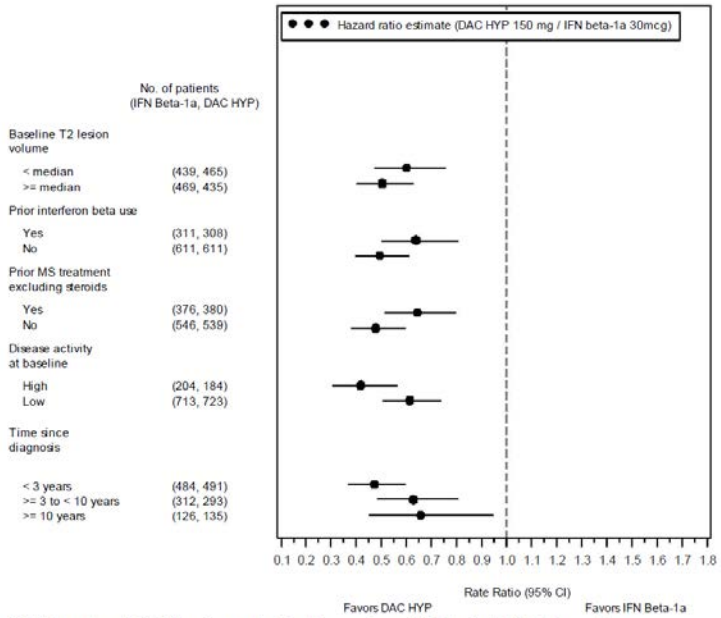
NOTE: Rate ratios and 95% CI based on negative binomial regression model adjusted for the following covariates (excluding covariates defining the subgroup): baseline relapse rate, history of prior IFN beta use, baseline EDSS (<=2.5 vs >2.5) and baseline age (<=35 vs >35).

Figure 12. Annualised relapse rate by demographic subgroups, Study 205MS301 (continued)



NOTE: Rate ratios and 95% CI based on negative binomial regression model adjusted for the following covariates (excluding covariates defining the subgroup): baseline relapse rate, history of prior IFN beta use, baseline EDSS (<=2.5 vs >2.5) and baseline age (<=35 vs >35).

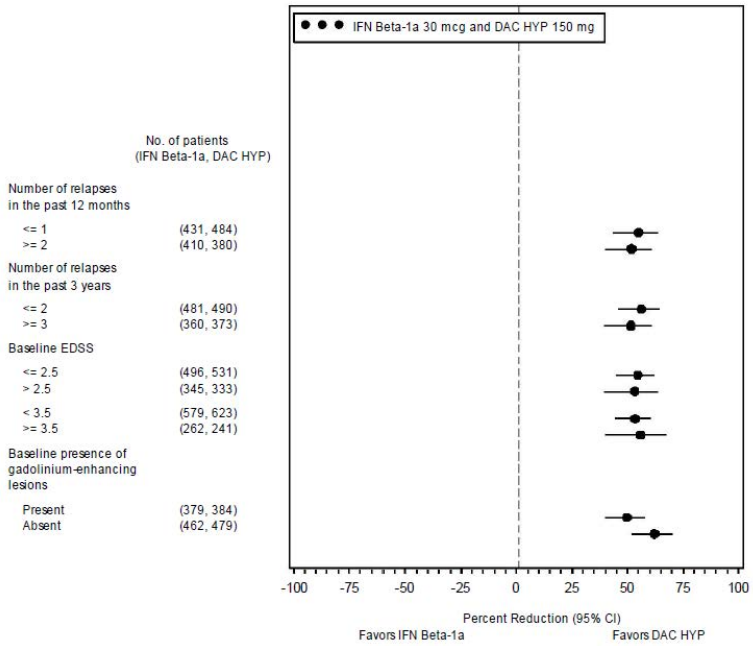
Figure 12. Annualised relapse rate by demographic subgroups, Study 205MS301 (continued)



NOTE: Rate ratios and 95% CI based on negative binomial regression model adjusted for the following covariates (excluding covariates defining the subgroup): baseline relapse rate, history of prior IFN beta use, baseline EDSS (<=2.5 vs >2.5) and baseline age (<=35 vs >35).

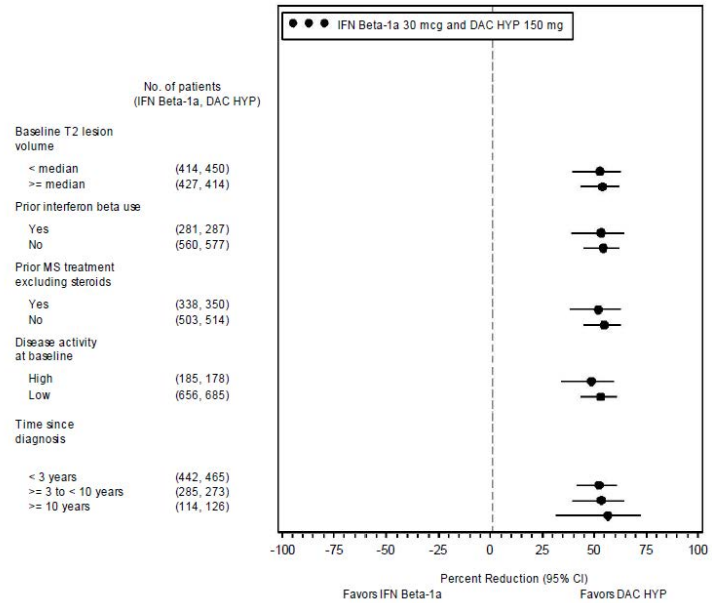
The sponsor also performed subgroup analyses for the main MRI endpoint, new or enlarging T2 lesions, and these consistently favoured DAC HYP, as shown in Figure 13 below. All subgroups in the table show a significant benefit for DAC HYP.

Figure 13. New or newly-enlarging T2 lesions by baseline disease characteristics



NOTE: Percent reduction and 95% CI based on negative binomial regression adjusted for the following covariates (excluding covariates defining the subgroup): baseline volume of T2 hyperintense lesions, history of prior IFN beta use and baseline age (<=35 vs >35).

Figure 13. New or newly-enlarging T2 lesions by baseline disease characteristics (continued)



NOTE: Percent reduction and 95% CI based on negative binomial regression adjusted for the following covariates (excluding covariates defining the subgroup): baseline volume of T2 hyperintense lesions, history of prior IFN beta use and baseline age (<=35 vs >35).

Subgroup analyses for the endpoint of sustained disability progression did not identify any subgroup in which the effects of DAC HYP and IFN β-1a were significantly different. This reflects the primary analysis of this endpoint in the full cohort, where a favourable trend was identified but no significant difference was observed.

Figure 14. Sustained disability progression (measured by increase in EDSS) by baseline disease characteristics

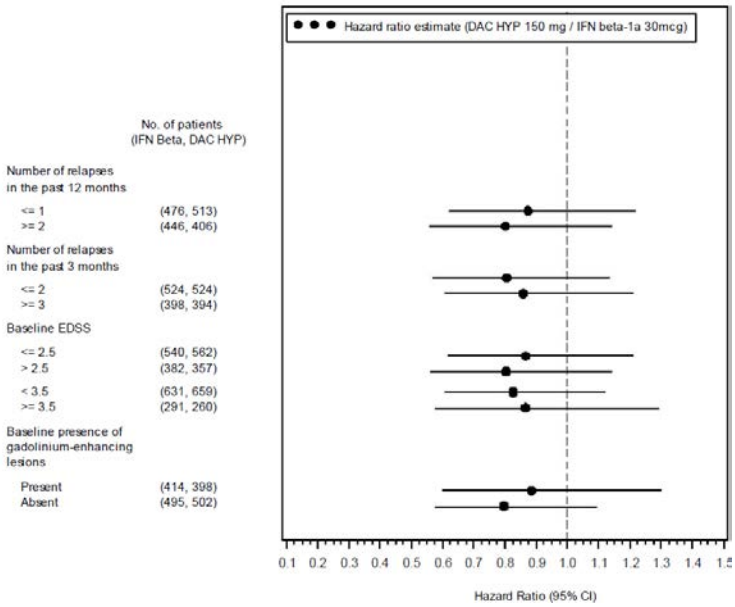
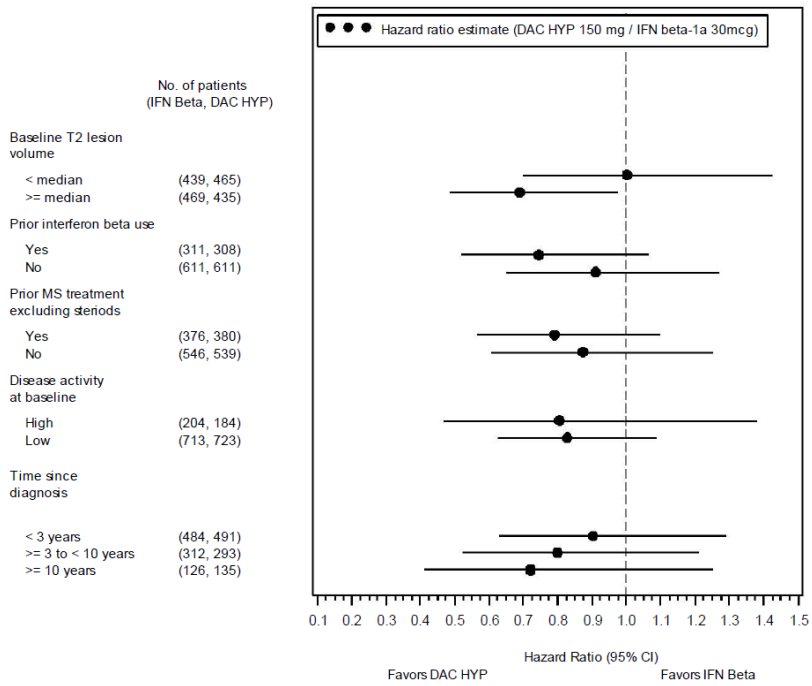


Figure 14. Sustained disability progression (measured by increase in EDSS) by baseline disease characteristics (continued)



NOTE: Hazard ratios and 95% CI based on Cox Proportional Hazards model, adjusted for the following covariates (excluding covariates defining the subgroup): baseline EDSS values as continuous variable, history of prior IFN-beta use and baseline age (<=35 vs >35).

Subgroup analysis for high disease activity versus low disease activity

A new subgroup analysis based on high disease activity versus low disease activity was suggested by the EMA, and is discussed in the following section concerning data supplied following EMA questions. Prior to receiving this request, the sponsor had already conducted their own subgroup analysis of high disease activity and low disease activity subgroups, and these results are included in the figures above. Although a similar analysis in Study 205MS201 had been post hoc, the definition of these subgroups in Study 205MS301 was prospective, and subjects were categorized as having low or high disease activity at baseline. High disease activity was defined as ≥ 2 relapses in the year prior to randomisation and ≥ 1 Gd lesion on the baseline MRI.

By this definition, there was a similar proportion of low disease activity subjects in both treatment groups, as shown in the figures above (IFN β-1a, 713/917, 77.8%; DAC HYP 723/907, 79.7%). The superiority over IFN β-1a was statistically significant in both subgroups (see Figure 7, above). There was a trend to DAC HYP producing a better reduction in ARR in the high-disease-activity subgroup than in the low-disease-activity subgroup.

For the secondary endpoint of new T2 lesions, a similar benefit was observed in both subgroups, and each subgroup achieved a statistically significant result showing superiority of DAC HYP relative to IFN β-1a.

For disease progression, a significant benefit with DAC HYP was not demonstrated for the cohort as a whole, or for any subgroup, including subgroups defined on the basis of disease activity.

7.2. Supplementary material submitted to the EMA in relation to pivotal efficacy data

7.2.1. EMA Question 70

> *'The indication claimed by the applicant is for the treatment of Relapsing Multiple Sclerosis. However in order to gain this unrestricted indication the applicant should further justify a positive effect on disability progression in relapsing forms of Secondary Progressive Multiple Sclerosis. The absence of such an effect may need to be reflected in the European Summary of Product Characteristics.'*

This question relates to an issue already flagged in this SCER: the sponsor performed two pivotal studies with entry criteria that explicitly excluded SPMS, but nonetheless proposes an indication that applies to the excluded MS category. The EMA question is specifically framed in terms of the endpoint of disability progression, but similar concerns could be raised about the overall efficacy of DAC HYP in SPMS, including its effect on relapse rate.

In their response to this question, the sponsor attempted to defend their extension of the indication beyond the original target population. The evaluator was not convinced that the arguments raised by the Sponsor were valid, and strongly recommends that the indication in the PI should match the entry criteria of the pivotal studies. At a minimum, this would mean that DAC HYP was restricted to subjects with RRMS. A strict interpretation of the pivotal studies would require that DAC HYP was further restricted to subjects with RRMS and evidence of recent disease activity, as evidenced by at least 2 relapses in the previous 3 years (or a radiological substitute for clinical relapses).

7.2.1.1. Sponsor's response

The sponsor began by pointing out that some subjects in the pivotal studies had SPMS, despite the exclusion criteria:

> *'In the clinical development of DAC HYP in MS, the 2 pivotal trials were of sufficient duration and size that subjects included in these trials could be identified as having SPMS with superimposed relapses based on the observation of sustained disability progression that occurred independently of, or in the absence of, clinical relapses.'*

The evaluator acknowledges that the pivotal studies were almost certainly contaminated with some subjects who had SPMS. Given that SPMS was explicitly listed as an exclusion criterion, this raises a number of difficulties in interpretation. If subjects had SPMS at baseline, then they were not entered into the study appropriately; conversely, if subjects developed SPMS during the study, then they were not prospectively randomised after entering the subgroup of interest.

For subjects who had SPMS at baseline, it might be possible to perform a post hoc subgroup analysis, if such subjects could be reliably identified, but their accidental inclusion does not constitute a clear prospective assessment of efficacy in SPMS, particularly if they only constituted a small proportion of the overall cohort. If they constituted a large proportion of the cohort, it implies widespread disregard for the entry criteria, and suggests the study was not methodologically sound.

For subjects who developed possible SPMS during the study, there are also difficulties with interpretation. Progression in the sense mentioned in the sponsor's statement above, 'sustained disability progression', could be attributed to two quite different processes: experiencing a relapse and failing to recover fully (giving a stepwise increase in disability), or progressing slowly in the absence of a relapse (giving a gradual increase in disability). Only the latter counts as SPMS, though it is difficult to separate the two categories accurately in clinical practice.

A third category also needs to be excluded, spurious progression, in which subjects have not yet had time to exhibit recovery from their most recent relapse; it is important that the methodological requirement for assessing disability progression at a time distant from the

relapse is not conflated with the issue of whether the increased disability was caused by a relapse. The sponsor's expression 'independently of, or in the absence of, relapses' and much of their subsequent discussion, is ambiguous about whether an aetiological or a temporal independence is intended. Only those subjects with gradual disease progression during the study could count as having developed SPMS post-baseline, and the number of such subjects is uncertain. The mere presence of subjects with 'confirmed disability progression' does not prove that all or most of those subjects had gradual disease progression, and therefore had SPMS (much less that they had SPMS at baseline).

If subjects developed SPMS during the pivotal studies, then this could be seen as a treatment failure and it is not logical to conclude that a benefit demonstrated for the whole cohort applies equally well to those who experienced a treatment failure. Even if it were possible to identify a subgroup in which true SPMS developed during the studies (by excluding subjects in whom the deterioration was causally associated with a relapse), it would be difficult to draw any conclusions from a subgroup analysis of their subsequent course. Firstly, their time left in the study would have been shortened by the requirement that they develop SPMS prior to entering the subgroup of interest, reducing the time left available in which they could reach new clinical endpoints. Secondly, as a subgroup, they are unlikely to have been present in large enough numbers for their assessment to achieve statistical power. Thirdly, their non-random entry into the SPMS subgroup while already on treatment would raise major issues of interpretation because treatment allocation would not necessarily be random across all subjects with SPMS. Fourthly, the fact that they had progressed despite taking DAC HYP is likely to mean that they are a subgroup in which DAC HYP efficacy was limited, so such a subgroup analysis could produce disappointing results.

The only analysis that would properly address the efficacy of DAC HYP in SPMS would be one that took subjects with clinically overt SPMS and examined them prospectively, in a randomised study. When dedicated SPMS studies have been attempted with other disease-modifying agents, the results have often been disappointing, and it is widely recognised that SPMS is a more difficult entity to treat than early, non-progressive RRMS. Decades of experience with MS research has shown that immune-modifying agents have blunted efficacy in this disease category, so efficacy in SPMS subjects must be demonstrated directly, not inferred from the results in RRMS subjects, or from the accidental inclusion of SPMS subjects in studies of RRMS.

The sponsor's response continued:

> '*Furthermore, analysis of these subjects provided evidence that DAC HYP was more effective than IFN β-1a at preventing the progression of sustained disability progression that occurred independently of clinical relapses.'*

This statement appears to be in direct contradiction of the fact that Study 205MS301 did not show a significant effect on sustained progression of disability for the cohort as a whole, or for any prospectively defined subgroup. A nominally significant result was shown for the tertiary endpoint of 24 week confirmed progression, but not for the key secondary endpoint of 12 week confirmed progression. In EMA Question 94 (discussed below), a subgroup was identified post hoc that a showed nominally significant benefit for DAC HYP on progession, but this subgroup analysis was not based on the presence of SPMS. From the quoted sentence above the expression 'analysis of these subjects' was used in the context of a discussion of SPMS. This implies that a subgroup analysis of SPMS subjects was undertaken and that this subgroup had reduced progression on DAC HYP. No prospective subgroup analysis addressed this issue, because SPMS subjects were not supposed to be enrolled. The statement therefore appears to be unjustified.

Both pivotal studies included some subjects with elevated EDSS at baseline, which could arguably serve as an imprecise surrogate marker for the presence or absence of SPMS. One section of the sponsor's response to EMA Question 70 proposed that EDSS ≥ 3.5 could be

interpreted this way. In the same section, the sponsor also appeared to suggest additional post hoc methods of identifying SPMS subjects, as follows:

> *'In order to explore the issue of DAC HYP's effect on patients with SPMS with superimposed relapses, the following analysis was performed for Study 301:*

1. *Identification of subjects with higher levels of baseline disability: EDSS ≥3.5, when the transition to SPMS is common.*
2. *Definition of a gradual worsening of neurologic function using standard criteria to measure disability progression in SPMS clinical trials that is based on a composite measure of 6-month confirmed worsening on either:*
   a. *EDSS,*
   b. *20% worsening in gait as measured by the T25FW, or*
   c. *20% worsening on upper extremity function as measured by the 9HPT. Progression is confirmed at a visit at least 6 months later and also at the last study visit.*
3. *Requirement that both the initiation of the neurologic worsening and the 6-month confirmation of the neurologic worsening occur in the absence of relapse within the prior month of the evaluation. As a further sensitivity analysis to ensure that new clinical relapses were not the cause of the worsening, an additional analysis of 6-month confirmed worsening was performed that was restricted to subjects who were relapse free during the entirety of Study 205MS301.'*

Each of these proposed methods for identifying SPMS subjects is potentially flawed, mostly because relapses in RRMS subjects can cause deficits in all of the domains studied: elevations of EDSS, worsening of gait, and worsening of upper limb function. Excluding a relapse in the previous month reduces the risk that a current relapse will be misinterpreted as disease progression, but does not prevent a previous relapse with incomplete recovery from being misidentified as a marker of SPMS. Note that the sponsor's original definition of confirmed progression required a 12-week delay to minimise the inclusion of cases with spurious progression, not the one-month delay proposed for this analysis. Also, note that the category of sustained progression is not limited to sustained gradual progression: sustained EDSS progression could include a stepwise deterioration due to a relapse, and would not identify a patient as having SPMS.

Overall, then, the sponsor's proposed subgroup analyses do not clearly identify a subgroup of subjects with SPMS. Even if they did, they would constitute a special subset of occult SPMS subjects – those who, at baseline, were explicitly thought by their clinicians not to have SPMS, but had it or developed it anyway. Even if efficacy were confirmed in such a subgroup, this would not apply to the broader disease category of overt SPMS. Further discussion of this important issue is found in the conclusions of this section.

Prior to submitting this supplementary material, the Sponsor had already performed a basic subgroup analysis based on EDSS. In Study 205MS201, subgroup analysis of subjects with higher baseline EDSS (> 2.5 versus ≤ 2.5) showed that, even in subjects with higher EDSS, DAC HYP has favourable effects on the primary endpoint of annualized relapse rate, compared to placebo. A similar subgroup analysis for the minor endpoint of disease progression appears not to have been performed in Study 205MS201, which is understandable as this was a tertiary endpoint and the study was too brief for a robust assessment of progression.

In Study 205MS301, subgroup analyses based on EDSS showed a benefit of DAC HYP over IFN β-1a on ARR in subjects with moderately elevated EDSS (> 2.5 versus ≤ 2.5) but the benefit was reduced, and not statistically significant, in subjects with even higher EDSS (≥ 3.5). For the secondary endpoint of disease progression, favourable trends were observed across a range of EDSS categories. The treatment effect for this endpoint was not statistically significant for the

cohort as a whole, so it is not surprising that EDSS subgroups also failed to show a significant reduction in disease progression.

Given that the higher EDSS subjects did not have overt SPMS at baseline (according to the protocol), the favourable trends for reduced ARR and the weaker trends for reduced progression cannot be generalised to other subjects with high EDSS and overt SPMS. The high-EDSS subgroups (> 2.5 or ≥ 3.5) would have included many subjects in whom baseline EDSS was elevated because of previous relapses with incomplete recovery, rather than because of SPMS. Indeed, according to the protocol, this was the only permissible reason for having a high baseline EDSS.

In the sponsor's response to EMA Question 70, the main efficacy measure of interest was disease progression, which was subjected to a new analysis, with alternate definitions of progression based on a composite of EDSS, timed walk and 9-hole peg test.

The sponsor's presentation of these supplementary analyses was somewhat unclear. Their discussion was primarily focussed on the extent to which DAC HYP was effective in subjects with SPMS that were inadvertently included in the pivotal studies or who developed SPMS during the studies. The proposed supplementary definitions of progression could be interpreted as an alternative means of identifying SPMS for potential subgroup analysis, and they were apparently introduced in that context. Further inspection of the tables suggests that these additional progression measures were used as alternative progression endpoints in the overall cohort, not as a means of identifying an SPMS subgroup. The sponsor was asked to confirm the nature of this analysis, with the following question:

> *'In the sponsor's response to EMA Question 70, two tables were supplied intended to address the efficacy of DAC HYP in subjects with Secondary Progressive MS (SPMS). From the sponsor's brief presentation of this data, it is somewhat unclear what analysis has actually been performed. Is the proportion with confirmed progression listed in each column of the table simply the proportion of the total EDSS-specific cohort that were considered to have progressed according to each of the listed criteria?*
>
> *For instance, in the first table, does the '0.241' listed opposite the Timed 25-Foot Walk under IFN β-1a simply mean that, of the 291 subjects at risk in this EDSS category, approximately 24.1% showed progression identified on the 25-Foot Walk? (Or similarlybut more precisely that the Proportional Hazards Model predicts that 24.1% would progress?)*
>
> *If so, is this exercise primarily an analysis of the risk of developing SPMS (identified by a range of markers), given a particular baseline EDSS? Was any analysis done of the risk of progression after being identified as having SPMS by any of these tests? If not, this exercise appears to consist of no more than a new subgroup analysis of EDSS categories, using a non-standard definition of progression instead of the protocol-specified definition as a 12-week sustained EDSS worsening. The analysis does not appear to provide any prospective assessment of subjects with SPMS.*

The core part of their answer is as follows:

> *'The values listed in the table are the estimated proportion of subjects who experienced disability progression by the listed criteria based on the Kaplan Meier product limit method (for example, from Table 1, 24.1% of IFN β-1a-treated subjects with baseline EDSS ≥ 3.5 are estimated to have 6-month confirmed 20% worsening on the timed 25-foot walk that did not start and was not confirmed at a visit within 29 days of an MS relapse).'*

This response confirms that the scores on the 9HPT and 25FW have been used as efficacy outcome variables, not as a means of identifying SPMS subjects for subsequent subgroup analysis. The relevance of this analysis to subjects with SPMS is indirect.

The results of the sponsor's reanalysis of progression are shown in the two tables below (Tables 14 and 15). The use of the term 'Independent of Relapse' in the title of the first table is

potentially misleading, because it merely indicates that no relapse had occurred in the month prior to the documentation of progression, not that the progression was causally independent of a relapse. In the second table, the analysis is restricted to subjects without any on-study relapses, and it appears more likely that, in such subjects, the observed progression does actually identify these subjects as having probable SPMS. Note that the number of such subjects was low: only 163 and 154 in the IFN β-1a and DAC HYP groups, respectively, had EDSS ≥ 3.5 and were free of relapses, and less than a quarter of these (23.4% in the IFN β-1a group) could be considered to have had SPMS by the 'Composite' progression measure at the end of the study. At baseline, none of them were thought to have SPMS by the enrolling clinician, and even if we accept that all such subjects had SPMS, it is unclear how many of them had SPMS at baseline and how many of them developed it during the study. In nearly all of the subgroups identified in this manner, and over a range of different definitions of progression, there are favourable trends for DAC HYP but the treatment benefit relative to IFN β-1a is not significant. Indeed, of 24 analyses across the two tables, only one had a 95% CI that excluded unity in favour of DAC HYP. Contrary to the sponsor' s conclusions, this analysis suggests that the number of SPMS subjects in the pivotal studies was very low, that the overall results cannot be extended to the SPMS population, and that the benefit of DAC HYP over IFN β-1a is not robust for the endpoint of progression.

Table 14. Summary of confirmed progression independent of relapse in Study 205MS301

Proportion subjects with confirmed progression at 144 weeks independent of relapse

EDSS range	Outcome	IFN beta-1a 30 mcg	DAC HYP 150 mg	HR (95% CI) (b)
>=3.5	Number of subjects evaluated	291	260	
	Composite	0.331	0.236	0.73 (0.51, 1.04)
	Timed 25-Foot Walk	0.241	0.153	0.66 (0.43, 1.01)
	Nine-Hole Peg (a)	0.078	0.070	0.92 (0.46, 1.83)
	EDSS	0.153	0.127	0.86 (0.52, 1.43)
>=4.0	Number of subjects evaluated	179	159	
	Composite	0.391	0.285	0.73 (0.48, 1.11)
	Timed 25-Foot Walk	0.275	0.182	0.66 (0.40, 1.10)
	Nine-Hole Peg (a)	0.101	0.085	0.79 (0.36, 1.75)
	EDSS	0.193	0.157	0.84 (0.47, 1.49)
>=4.5	Number of subjects evaluated	97	84	
	Composite	0.445	0.344	0.77 (0.44, 1.33)
	Timed 25-Foot Walk	0.297	0.173	0.58 (0.29, 1.15)
	Nine-Hole Peg (a)	0.094	0.085	0.85 (0.28, 2.54)
	EDSS	0.281	0.237	0.91 (0.47, 1.76)

Note: Estimated proportion of subjects with confirmation is based on the Kaplan Meier product limit method.
(a) Analysis excludes subjects with missing baseline data for Nine-Hole Peg Test.
(b) Based on Cox Proportional Hazards model, adjusted by baseline value of the corresponding MSFC component or EDSS, history of prior IFN beta use, and baseline age (age <= 35 vs age >35). Analysis on composite adjusted for baseline EDSS, baseline Timed 25-Foot Walk Test, baseline Nine-Hole Peg Test, history of prior IFN beta use, and baseline age (age <= 35 vs age >35).

Table 15. Summary of confirmed progression in relapse-free population in Study 205MS301

Proportion subjects with confirmed progression at 144 weeks and relapse free

EDSS range	Outcome	IFN beta-1a 30 mcg	DAC HYP 150 mg	HR (95% CI) (b)
>=3.5	Number of subjects evaluated	163	154	
	Composite	0.234	0.143	0.67 (0.36, 1.22)
	Timed 25-Foot Walk	0.164	0.091	0.49 (0.23, 1.03)
	Nine-Hole Peg (a)	0.064	0.031	0.50 (0.15, 1.67)
	EDSS	0.087	0.072	1.16 (0.46, 2.93)
>=4.0	Number of subjects evaluated	101	88	
	Composite	0.284	0.134	0.50 (0.22, 1.15)
	Timed 25-Foot Walk	0.189	0.074	0.34 (0.12, 0.97)
	Nine-Hole Peg (a)	0.066	0.000	NA
	EDSS	0.153	0.092	0.76 (0.27, 2.18)
>=4.5	Number of subjects evaluated	56	45	
	Composite	0.392	0.172	0.49 (0.17, 1.39)
	Timed 25-Foot Walk	0.221	0.056	0.23 (0.05, 1.13)
	Nine-Hole Peg (a)	0.095	0.000	NA
	EDSS	0.290	0.145	0.62 (0.20, 1.98)

Note: Estimated proportion of subjects with confirmation is based on the Kaplan Meier product limit method.
(a) Analysis excludes subjects with missing baseline data for Nine-Hole Peg Test.
(b) Based on Cox Proportional Hazards model, adjusted by baseline value of the corresponding MSFC component or EDSS, history of prior IFN beta use, and baseline age (age <= 35 vs age >35). Analysis on composite adjusted for baseline EDSS, baseline Timed 25-Foot Walk Test, baseline Nine-Hole Peg Test, history of prior IFN beta use, and baseline age (age <= 35 vs age >35).

7.2.1.2. Evaluator's conclusion

Overall, despite the sponsor's arguments to the contrary, there has been no clear assessment of the efficacy of DAC HYP in subjects with SPMS, given that the pivotal studies explicitly excluded this disease category, efficacy in this population is unknown.

Subjects with a moderately high baseline EDSS (attributed by their clinician to relapses rather than to SPMS) appeared to respond to DAC HYP with a reduced relapse rate in both pivotal studies. This implies that the drug has efficacy in subjects with high EDSS when it is due to incomplete recovery from previous relapses, but it does not allow a generalisation to subjects with a high EDSS due to SPMS, which was explicitly not intended to be assessed in the pivotal studies.

The data for progression are less favourable than those for ARR. According to prospective analysis methods in Study 205MS201, the study did not show a significantly reduced rate of progression relative to placebo. The proposed DAC HYP dose of 150 mg achieved nominal statistical significance for progression in this study, but only if issues with multiplicity are ignored. (Relative to placebo, the hazard ratio for disability progression was 0.43 (95% CI: 0.21 to 0.88) in the DAC HYP 150 mg group and 0.57 (95% CI: 0.30 to 1.09) in the DAC HYP 300 mg group.) No subgroup analysis of progression data in SPMS subjects from Study 205MS201 was presented.

In Study 205MS301, high-EDSS subjects did not have a significantly reduced rate of progression relative to IFN β-1a, using the major prospective definition of progression (12 week sustained EDSS worsening) in keeping with the lack of a significant effect of DAC HYP on progression in the cohort as a whole.

The sponsor's supplementary analyses of Study 205MS301, shown in the tables above, appear to identify some subjects who had probable SPMS at the end of the study, but it remains completely unknown how many of these subjects had SPMS at the time of study entry. The group of subjects with SPMS at baseline is likely to have been too small to allow any robust inferences to be drawn, even if such subjects could be identified. At best, the provided tables indicate that DAC HYP may reduce the development of SPMS in high EDSS subjects, as suggested by favourable trends in this post hoc analysis, but this is not a statistically robust finding and would require confirmation in a prospective study. The tables do not constitute an analysis of DAC HYP efficacy in subjects with SPMS.

Even if it were known that DAC HYP reduced the development of SPMS, this would not establish that it has efficacy in SPMS, and it is unfortunate that the Sponsor's discussion frequently conflated these issues. If we translate the issue to a different domain, it becomes clear that the ability to prevent a disease is not the same as efficacy in treating the disease (showing that a vaccine could prevent influenza would not logically imply that it also had efficacy in the treatment of influenza.) For SPMS, the situation is more complex, because there is a spectrum of disease between MS dominated by relapses and MS dominated by progression. There are a priori reasons to suspect that mechanisms of action that prevented relapses in early disease might still be beneficial when MS has reached a secondary progressive phase. Nonetheless, there is also substantial evidence from decades of MS research suggesting that immune therapies have less efficacy in SPMS, and this means that efficacy in SPMS must be demonstrated explicitly. If this were not the case, the sponsor would not have listed SPMS as an exclusion criterion in the first place.

To some extent, the sponsor's entire discussion in their response to EMA Question 70 is based on the assumption that accidental violations of inclusion criteria justify subsequent broadening of the target population. This reasoning is rejected for three main reasons:

1. It encourages a second round of 'bracket creep' in the attempt to identify the target population. Many of the standard MS categories lack clear boundaries, including the category of SPMS. It is therefore inevitable that a large MS study would inadvertently

include some subjects in whom the categorisation of their MS was open to debate and potentially subject to redefinition. Importantly, though, the same is also true of clinical characterisation of MS when prescribing a drug in an MS patient. Some degree of blurring of the boundaries is inevitable in both trial recruitment and clinical prescribing and in the absence of further evidence we should assume that the blurring is about the same in both situations. (It is performed by the same clinicians treating the same disease). If some subjects in the pivotal studies had a clinical appearance suggestive of RRMS but actually had occult SPMS and if those subjects appeared to benefit from DAC HYP, then the same is likely to be true of subjects in clinical practice that appear to have RRMS and actually have occult SPMS. Such subjects would end up treated anyway, using an indication that matched the entry criteria. By matching the target population to the study entry criteria, the blurring of the diagnostic categories in both situations would be expected to include the same subjects, and produce similar benefits. Conversely, if the imprecision of the diagnostic categories in a study is used to justify a more inclusive indication, then the imprecision of the clinical categorisation at the time of prescribing adds a second round of blurring, leading to inclusion of subjects not represented in the original study. In this particular context, if the indication for DAC HYP were worded as 'relapsing forms of MS' subjects with overt SPMS who were not actually experiencing ongoing true relapses could be included because:

a. their MS had begun with relapses, and therefore they could be considered to have a relapsing form of MS;

b. the tendency of MS-related pre-existing deficits to flare in the context of inter-current illnesses ('pseudo-relapses') could be construed as ongoing relapses;

c. the tendency of gradual functional decline to cross non-gradual milestones could produce the appearance of a stepwise decline and be construed as a relapse, such as first need of a walking aid, or first use of a wheelchair;

d. intercurrent soft-tissue injuries, which are common complications of MS-related motor disability, could produce temporary deteriorations and be classified as relapses;

e. non-MS-related deficits, such as complicated migraines or middle ear infections, could be labelled as relapses;

f. clinicians or patients could misrepresent the underlying clinical symptoms deliberately or subconsciously because of their hope that 'doing something' for the MS is better than giving up.

For all of these reasons, some spread of the target population could be expected at the time of prescribing, which should not be added to over-inclusiveness in the wording of the indication.

2. The accidental inclusion of some subjects with occult SPMS in the pivotal studies does not prove that the drug works in this category. Studies only provide clear evidence of efficacy if they are prospective tests of well-defined hypotheses, and accidental inclusion of subjects with a different condition constitutes a methodological flaw. If some subjects with small-vessel cerebrovascular disease had been misdiagnosed as having MS and been accidentally recruited, their inclusion would clearly not support the claim that DAC HYP worked in small-vessel disease.

3. The sponsor's original study design with its explicit entry criteria was a tacit admission that these subjects are more difficult to treat than RRMS subjects, or at least different to RRMS subjects. The sponsor is yet to account for the inherent contradiction in their approach: their study design explicitly excluded these difficult-to-treat subjects, but they have nonetheless proposed including them in the indication.

As there is a spectrum of MS disease types from pure RRMS to pure non-relapsing SPMS, it is likely that some efficacy is achievable in subjects with SPMS, at least for those subjects still experiencing relapses, but the onus is on the sponsor to asses that potential for efficacy, rather than inferring it from studies that explicitly excluded such subjects. The indication in the PI should therefore be changed to match the entry criteria of the pivotal studies. There is, ultimately, no logical rationale for performing a study with one set of entry criteria and then proposing that clinicians prescribe on the basis of more inclusive criteria.

7.2.2. EMA Question 94

> *'The applicant is invited to redefine 'highly active MS' as per Tecfidera and Aubagio SmPC definitions, and should perform a comparison of safety data (for) high versus low disease activity, as this will have impact in the benefit/risk discussion.'*

This question, as summarised in the material provided to the clinical evaluator, did not explicitly ask for an efficacy analysis using the new definition, but this would appear to the main point of the redefinition, and it is implied that the safety analysis was additional. As discussed below, the estimation of the efficacy of DAC HYP was potentially affected by the redefinition, but the safety assessment was unaffected. Discussion of safety results in this subgroup analysis is included in the Safety section of this report.

Table 16. Original and new sponsor definition of 'highly active MS'

Original definition Study 205MS201	Original definition Study 205MS301	New definition
As a post-hoc analysis, the efficacy of DAC HYP was also evaluated in subjects with high disease activity at baseline, defined as: ≥ 2 relapses in the year prior to randomisation and; ≥ 1 Gd-enhancing lesion at baseline as well in subjects with and without prior MS treatment experience (excluding steroids). Post hoc	High disease activity was defined as: ≥ 2 relapses in the year prior to randomisation and; ≥1 Gd+ lesion on the baseline MRI. Prospective	Definition 1: Subjects with ≥ 2 relapses in 1 year and; ≥ 1 Gd-enhanching lesions on brain MRI OR Definition 2: Subjects who failed to respond to a full and adequate course (≥ 1 year) of β-IFN, having had ≥ 1 relapse in the previous year while on therapy, and; ≥ 9 T2 lesions in cranial MRI or ≥ 1 Gd-enhancing lesion, or having an unchanged or increased relapse rate in the prior year as compared to the previous 2 years Post hoc

It should be noted that, for each pivotal study, the sponsor had already performed their own subgroup analysis of subjects with high disease activity.

In Study 205MS201, the original analysis suggested that efficacy was similar in the high-disease-activity and low-disease-activity subgroups for most major endpoints including the primary endpoint, ARR, as well as radiological endpoints and the tertiary endpoint of disease

progression. Unfortunately, given that this was a post hoc analysis, it lacks formal statistical validity. The EMA definition is also post hoc, but because this definition has been used for other products, it could potentially be considered at least partly independent of the results actually obtained with DAC HYP.

In Study 205MS301, the original definition of high disease activity appeared to be prospective. The original subgroup analysis of high disease activity and low disease activity suggested that, for the primary endpoint of ARR, superiority of DAC HYP over IFN β-1a was statistically significant in both subgroups. Similarly, for new T2 lesions, a similar and statistically significant benefit of DAC HYP relative to IFN β-1a was observed in both subgroups. For disease progression, a significant benefit with DAC HYP relative to IFN β-1a was not demonstrated for the cohort as a whole, or for any subgroup, including subgroups defined on the basis of disease activity.

The main difference between the original definition of high disease activity and the new definition proposed by the EMA is that the new definition allows subjects to be defined as having high disease activity on the basis of having failed treatment with beta interferons. Thus, it does not simply identify a group with high disease activity, but high activity and/or interferon resistance. If the results of subgroup analyses are different with this definition, then the difference is likely to be due to the addition of interferon-resistant subjects to the subgroup. The revised definition and the subsequent selective enrichment of this subgroup with interferon-resistant subjects would be expected to improve the apparent efficacy of DAC HYP relative to an interferon control therapy. As shown in the sponsor's response, below, this is what was observed. Conversely, subjects with low disease activity did not show a significant benefit with DAC HYP relative to IFN β-1a.

The fact that the proposed definition was post hoc further undermines the statistical validity of the analysis. The p-values cited have been produced by the application of statistical tests that assume, incorrectly in this instance, that they will be applied as isolated tests of a well-defined, prospective hypothesis.

The sponsor's response consisted of two and half pages of text, followed by several pages of tables. The text is reproduced in its entirety below, followed by a discussion that is limited to the underlined and bolded efficacy endpoints. The safety endpoints mentioned in the response are discussed separately, in the Safety section of this report but there was no evidence of a substantially different safety profile in subjects with high or low disease activity.

7.2.2.1. Sponsor's response

The applicant redefined high disease activity as per the Tecfidera and Aubagio SmPCs. This modified definition added a second criterion to the definition used in the applicant's primary analysis as shown below:

1. Subjects with 2 or more relapses in 1 year, and with 1 or more Gd-enhancing lesions on brain MRI, or
2. Subjects who failed to respond to a full and adequate course (at least 1 year of treatment) of beta-interferon, having had at least 1 relapse in the previous year while on therapy, and at least 9 T2-hyperintense lesions in cranial MRI or at least 1 Gd-enhancing lesion, or having an unchanged or increased relapse rate in the prior year as compared to the previous 2 years

Although the definition above matches the definitions in the Tecfidera SmPC and Aubagio SmPC, the analysis in the Aubagio SmPC was based on the first criterion only as no data were available for the second criterion.

Subjects who did not meet the criteria for high disease activity were classified in our analyses as having low/unknown disease activity.

To facilitate the assessment of benefit/risk based on this new definition of high disease activity, analyses were performed on the data from Study 205MS201 and Study 205MS301 for the following endpoints by baseline disease activity level:

- Summary of AEs
- Incidence of maximum values in liver function tests (Study 205MS301 only)
- Annualised relapse rate (using INEC confirmed relapses)
- Number of new or newly enlarging T2 lesions
- 6-month sustained disability progression

Study 205MS201

In Study 205MS201, the overall AE profile was similar for the subjects with high and low/unknown disease activity at baseline. The incidence of AEs and SAEs reported were also similar among subjects with high disease activity and low/unknown disease activity. Notably the incidence of AEs in the high and low disease activity subgroups of the total DAC HYP group was similar for events in the Infections and Infestations SOC (53% and 52%, respectively) and the Skin and Subcutaneous Tissue Disorders SOC (16% and 21%, respectively). The results of the analyses of annualised relapse rate and new or newly enlarging T2 lesions by baseline disease activity demonstrate the superiority of DAC HYP over placebo for both the high and low/unknown disease activity subgroups. The reductions in the annualised relapse rate in the DAC HYP 150 mg group relative to placebo were similar, with a 52% reduction ($p = 0.0493$) in the high disease activity group and a 54% reduction ($p = 0.0003$) in the low/unknown disease activity (see Table 17, below) In the analysis of new or newly enlarging T2 lesions (see Table 18, below), the reduction relative to placebo was greater in the high disease activity group (78%, $p < 0.0001$) than in the low/unknown disease activity group (66%, $p < 0.0001$).

In the analyses of disability progression (see Table 19, below), treatment with DAC HYP 150 mg was associated with a markedly lower rate of 6-month sustained progression compared to placebo in both the high disease activity group (hazard ratio = 0.23, $p = 0.2034$) and the low/unknown disease activity group (hazard ratio = 0.24, $p = 0.0093$).

Study 205MS301

As was the case in Study 205MS201, there were no notable imbalances in the safety data between the high and low/unknown disease activity groups in Study 205MS301. The incidence of SAEs was greater in subjects with high disease activity as compared to subjects with low disease activity in both treatment groups, suggesting the differences were associated with baseline disease severity and were not indicative of treatment-related differences. In the DAC HYP arm, the incidence of AEs was slightly higher in the high disease activity subgroup as compared to the low disease activity subgroup for the Infections and Infestations SOC (70% versus 62%) and the Skin and Subcutaneous Disorders SOC (41% versus 62%). However, a similar trend was also seen in the IFN β-1a group, which suggests the differences are primarily a function of greater disease severity in these subjects.

Maximum values for liver function tests were also similar in the high and low/unknown disease activity groups of Study 205MS301. Most subjects in both subgroups had maximum values that were between ≤ 3 x ULN. The incidence of maximum values ≥ 5 x ULN was low and similar between the disease activity subgroups and the DAC HYP and IFN β-1a arms. The results of the analyses of annualized relapse rate and new or newly enlarging T2 lesions by baseline disease activity demonstrate the superiority of DAC HYP over IFN β-1a for both the high and low/unknown disease activity subgroups, with highly significant p values (< 0.0001). For annualised relapse rate (see Table 20, below), the effect relative to IFN β-1a was greater in the high disease activity group (rate ratio 0.497: 95% CI 0.397 to 0.621) than in the low/unknown disease activity group (rate ratio = 0.614: 95% CI 0.490 to 0.770). For new or newly enlarging

T2 lesions (see Table 21, below), the results by baseline activity were comparable (reductions of 53.7% and 52.3%, respectively, for high and low/unknown disease activity).

In Study 205MS301, there was a 43% reduction in 6-month sustained disability progression with DAC HYP compared to IFN β-1a in the high disease activity subgroup (hazard ratio = 0.57, p = 0.0102). No significant difference was evident between treatment groups in the low/unknown disease activity group (hazard ratio = 0.89, p = 0.5662) (see Table 22, below). The stronger treatment effect in the high disease activity subgroup may be due to a higher rate of disease progression in the IFN β-1a group, which provides more power to detect a treatment benefit. Conversely, the low rate of disease progression in the IFN β-1a arm provides less power to detect a treatment effect in the low disease activity subgroup. A similar pattern has been seen in other MS development programs in which a significant treatment benefit over IFN β-1a has been difficult to establish when there is a low progression rate. Nevertheless, the clearly superior findings of efficacy against disability progression compared to placebo in the low disease activity subgroup of Study 205MS201 provide evidence that DAC HYP does have a beneficial effect on disability progression in these subjects.

The results of these analyses demonstrate that the benefit/risk profile of DAC HYP remains favourable when high disease activity is redefined based on the Tecfidera/Aubagio SmPCs. The overall safety profile of DAC HYP is consistent in subjects with low and high disease activity at baseline in both studies. Likewise, DAC HYP provides a meaningful and consistent efficacy benefit over placebo and IFN β-1a whether measured in terms of relapses (annualized relapse rate), number of new/newly enlarging T2 lesions, or disability progression in subjects with both high and low disease activity at baseline. The differences between subgroups for some of the safety and efficacy results in both studies were generally observed in both the DAC HYP and control groups and were consistent with the greater level of disease activity at baseline.

The key tables mentioned in this response that are relevant to efficacy are reproduced below (with different numbering than in the sponsor's original text). The evaluator's interpretation of the significance of these results is included below the tables.

Table 17. Annualised relapse rate by disease activity and treatment, Study 205MS201

Annualized relapse rate by high disease activity subgroups, 205MS201

	Placebo	150 mg DAC HYP	300 mg DAC HYP
Baseline Disease Activity			
High			
n	37	46	38
Adjusted relapse rate (95% CI)	0.54 (0.344, 0.854)	0.26 (0.147, 0.463)	0.29 (0.159, 0.538)
Percent reduction (95% CI)		51.83 (-0.099, 76.820)	46.02 (-15.304, 74.726)
p-value vs placebo		0.0493	0.1152
Low/Unknown			
n	159	155	165
Adjusted relapse rate (95% CI)	0.44 (0.345, 0.559)	0.20 (0.139, 0.287)	0.21 (0.154, 0.301)
Percent reduction (95% CI)		54.47 (29.746, 70.491)	51.05 (26.163, 67.552)
p-value vs placebo		0.0003	0.0005

NOTE 1: High disease activity at baseline is defined as 1) patients with >= 2 relapses in the previous year and >= 1 Gd lesion at baseline, or 2) patients who have failed to respond to at least one year of treatment of beta-interferon, having had >= 1 relapse in the previous year while on therapy, and >= 9 T2 lesions or >=1 Gd lesion, or patients having an unchanged or increased relapse rate in the prior year as compared to the previous 2 years. Otherwise disease activity is 'Low/Unknown'.
2: Estimated from a negative binomial regression model adjusted for baseline EDSS (<=2.5 vs. >2.5), baseline age (<=35 vs. >35) and number of relapses in the 1 year prior to study entry.

Table 18. New or newly enlarged T2 lesions by disease activity and treatment, Study 205MS201

Number of new or newly-enlarging T2 lesions at Week 52 by high disease activity subgroups, 205MS201

	Placebo	150 mg DAC HYP	300 mg DAC HYP
Baseline Disease Activity			
High			
n	37	46	38
Adjusted mean number (95% CI)	15.66 (10.56, 23.23)	3.49 (2.413, 5.059)	3.17 (2.113, 4.762)
Percent reduction (95% CI)		77.69 (61.05, 87.23)	79.75 (63.97, 88.62)
p-value vs placebo		<0.0001	<0.0001
Low/Unknown			
n	158	153	162
Adjusted mean number (95% CI)	6.39 (5.069, 8.047)	2.18 (1.692, 2.802)	1.47 (1.136, 1.893)
Percent reduction (95% CI)		65.91 (51.87, 75.85)	77.04 (67.59, 83.74)
p-value vs placebo		<0.0001	<0.0001

NOTE 1: High disease activity at baseline is defined as 1) patients with >= 2 relapses in the previous year and >= 1 Gd lesion at baseline, or 2) patients who have failed to respond to at least one year of treatment of beta-interferon, having had >= 1 relapse in the previous year while on therapy, and >= 9 T2 lesions or >=1 Gd lesion, or patients having an unchanged or increased relapse rate in the prior year as compared to the previous 2 years. Otherwise disease activity is 'Low/Unknown'.
2: Estimated from a negative binomial regression model adjusted for baseline number of T2 lesions.

Table 19. Time to sustained progression by disease activity and treatment, Study 205MS201

	Placebo	150 mg DAC HYP	300 mg DAC HYP
High disease activity at baseline			
Number of subjects in ITT population	37 (100)	46 (100)	38 (100)
Number of subjects who progressed	4 (11)	1 (2)	0
Time (wk) to progression (a)			
25th percentile	NA	NA	NA
50th percentile	NA	NA	NA
Estimated proportion of subjects with progression at 52 weeks (a)	0.114	0.023	0.000
Hazard ratio and 95% CI (b)		0.23 (0.02-2.19)	ND
p-value vs placebo (b)		0.2034	ND
Low/Unknown disease activity at baseline			
Number of subjects in ITT population	159 (100)	155 (100)	165 (100)
Number of subjects who progressed	17 (11)	4 (3)	13 (8)
Time (wk) to progression (a)			
25th percentile	NA	NA	NA
50th percentile	NA	NA	NA
Estimated proportion of subjects with progression at 52 weeks (a)	0.111	0.027	0.084
Hazard ratio and 95% CI (b)		0.24 (0.08-0.70)	0.71 (0.34-1.47)
p-value vs placebo (b)		0.0093	0.3549

NOTE: High disease activity at baseline is defined as 1) patients with >= 2 relapses in the previous year and >= 1 Gd lesion at baseline, or 2) patients who have failed to respond to at least one year of treatment of beta-interferon, having had >= 1 relapse in the previous year while on therapy, and >= 9 T2 lesions or >=1 Gd lesion, or patients having an unchanged or increased relapse rate in the prior year as compared to the previous 2 years. Otherwise disease activity is 'Low/Unknown'.
(a) Estimated time to progression and proportion of subjects with progression based on the Kaplan-Meier product limit method.
(b) Hazard ratio and p-value estimated from a Cox proportional hazards model. Covariates included were baseline EDSS (<=2.5 vs >2.5), and age (<=35 vs >35). 'ND': Not Done due to no events.

Table 20. Annualised relapse rate by disease activity and treatment, Study 205MS301

Annualized relapse rate (INEC confirmed relapses) by high disease activity subgroups, 205MS301

	IFN beta-1a 30 mcg		DAC HYP 150 mg			
	n	Adjusted mean and 95% CI (a)	n	Adjusted mean and 95% CI (a)	Rate ratio and 95% CI (a)	P-value (a)
Disease activity at baseline						
High	386	0.530 (0.457, 0.614)	358	0.263 (0.220, 0.314)	0.497 (0.397, 0.621)	<0.0001
Low/Unknown	536	0.319 (0.266, 0.384)	561	0.196 (0.162, 0.237)	0.614 (0.490, 0.770)	<0.0001

NOTE 1: Only relapses confirmed by INEC are included in this analysis.
2: Data after subjects switched to alternative MS medications are excluded.
3: High disease activity at baseline is defined as 1) patients with >= 2 relapses in the previous year and >= 1 Gd lesion at baseline, or 2) patients who have failed to respond to at least one year of treatment of beta-interferon, having had >= 1 relapse in the previous year while on therapy, and >= 9 T2 lesions or >=1 Gd lesion, or patients having an unchanged or increased relapse rate in the prior year as compared to the previous 2 years. Otherwise disease activity is 'Low/Unknown'.
(a) Assessing the difference between the treatment groups, based on negative binomial regression model adjusted for the following covariates: baseline relapse rate, history of prior IFN beta use, baseline EDSS (<=2.5 vs >2.5) and baseline age (<=35 vs >35).

Table 21. New or newly enlarged T2 lesions by disease activity and treatment, Study 201MS301

Number of new or newly-enlarging T2 lesions at Week 96 by high disease activity subgroups, 205MS301

	IFN beta-1a 30 mcg		DAC HYP 150 mg			
	n	Adjusted mean and 95% CI (a)	n	Adjusted mean and 95% CI (a)	Percent reduction and 95% CI (a)	P-value (a)
Disease activity at baseline						
High	348	13.40 (11.52, 15.59)	339	6.21 (5.31, 7.26)	53.7 (42.4, 62.7)	<0.0001
Low/Unknown	493	6.80 (5.63, 8.21)	525	3.24 (2.72, 3.86)	52.3 (41.3, 61.3)	<0.0001

NOTE 1: Observed data after subjects switched to alternative MS medications are excluded. Missing data are not imputed. Only observed new or newly enlarging T2 lesions at the last visit of the subject up to Week 96 visit is used in this analysis.
2: High disease activity at baseline is defined as 1) patients with >= 2 relapses in the previous year and >= 1 Gd lesion at baseline, or 2) patients who have failed to respond to at least one year of treatment of beta-interferon, having had >= 1 relapse in the previous year while on therapy, and >= 9 T2 lesions or >=1 Gd lesion, or patients having an unchanged or increased relapse rate in the prior year as compared to the previous 2 years. Otherwise disease activity is 'Low/Unknown'.
(a) Estimated from a negative binomial regression model, adjusted for the following covariates: baseline volume of T2 hyperintense lesions, history of prior IFN beta use and baseline age (<=35 vs >35). To account for the timing of the MRI measurement, the logarithmic transformation of the scan number of the MRI assessment will be included in the model as the 'offset' parameter.

Table 22. Time to sustained progression by disease activity and treatment, Study 205MS301

Summary of time to 6-month sustained disability progression measured by increase in EDSS with multiple imputation by high disease activity subgroup, 205MS301

	IFN beta-1a 30 mcg		DAC HYP 150 mg			
	n	Estimated proportion of subjects with progression at 144 weeks (a)	n	Estimated proportion of subjects with progression at 144 weeks (a)	Hazard ratio and 95% CI (b)	P-value (b)
Disease activity at baseline						
High	386	0.240	358	0.122	0.57 (0.37, 0.87)	0.0102
Low/Unknown	536	0.143	561	0.130	0.89 (0.60, 1.32)	0.5662

NOTE: High disease activity at baseline is defined as 1) patients with >= 2 relapses in the previous year and >= 1 Gd lesion at baseline, or 2) patients who have failed to respond to at least one year of treatment of beta-interferon, having had >= 1 relapse in the previous year while on therapy, and >= 9 T2 lesions or >=1 Gd lesion, or patients having an unchanged or increased relapse rate in the prior year as compared to the previous 2 years. Otherwise disease activity is 'Low/Unknown'.
(a) Estimated time to progression and proportion of subjects with progression based on the Kaplan-Meier product limit method.
(b) Hazard ratio and p-value assessing the difference between the treatment groups were estimated from a Cox proportional hazards model. Covariates included were baseline EDSS (<=2.5 versus >2.5), and age (<= 35 versus >35).

7.2.2.2. Evaluator's conclusions

Most of the results in this supplementary analysis are concordant with the original analysis of the each study, and therefore do not provide any major new insights. To the extent that the results differ may be attributed to post hoc changes in the definition of progression (switching

from 12 week confirmed progression to 6 month confirmed progression) as well as a post hoc revision of the definition of high disease activity (including subjects with proven poor response to the active comparator of Study 205MS301).

For Study 205MS201, a significant benefit of DAC HYP over placebo was demonstrated for ARR in the 150 mg dose group for subjects with high and low disease activity, but for the 300 mg dose group a significant benefit was only demonstrated for low disease activity. The hazard ratios were similar for subjects with high disease activity administered DAC HYP 300 mg, but the analysis was underpowered. For T2 lesions, all four combinations of dose and disease activity produced significant superiority of DAC HYP over placebo. For disease progression, a significant benefit was only observed for one of the four combinations (DAC HYP 150 mg and low disease activity), but the general trends were favourable across doses and subgroups. A p-value was not calculated for the combination of DAC HYP 300 mg and high disease activity, because none of these subjects progressed. In general, the number of subjects with high disease activity was too low for robust statistical analysis., but the trends suggested that the relative benefit of DAC HYP over IFN β-1a was broadly similar across subgroups.

For Study 205MS301, this new subgroup analysis showed broadly similar benefits in both subgroups, when ARR and T2 lesion load were analysed ($p < 0.0001$ for each endpoint for each subgroup). This is broadly consistent with the original subgroup analysis. In both the original analysis and the supplementary analysis, the rate ratio for ARR was lower (more favourable) in the high-activity subgroup and higher (but still significantly in favour of DAC HYP) in the low-activity group.

Table 23. Table excerpt of ARR by baseline disease activity, Study 205MS301

	IFN beta-1a 30 mcg	DAC HYP 150 mg
Disease activity at baseline		
High (>= 2 relapses in the year prior to randomization and >= 1 Gd lesion at baseline MRI)		
n	204 (22)	184 (20)
Adjusted annualized relapse rate	0.677	0.282
and 95% CI (a)	(0.550, 0.834)	(0.221, 0.360)
Rate ratio (DAC HYP/IFN beta-1a)		0.417
and 95% CI (a)		(0.306, 0.568)
p-value vs IFN beta-1a (a)		<0.0001
Low		
n	713 (77)	723 (79)
Adjusted annualized relapse rate	0.322	0.197
and 95% CI (a)	(0.284, 0.365)	(0.171, 0.228)
Rate ratio (DAC HYP/IFN beta-1a)		0.613
and 95% CI (a)		(0.509, 0.739)
p-value vs IFN beta-1a (a)		<0.0001

NOTE 1: Only relapses confirmed by INEC are included in this analysis.
2: Data after subjects switched to alternative MS medications are excluded.
3: Numbers in parentheses are percentages.
(a) Assessing the difference between the treatment groups, based on negative binomial regression model adjusted for the following covariates (excluding covariates defining the subgroup): baseline relapse rate, history of prior IFN beta use, baseline EDSS (<=2.5 vs >2.5) and baseline age (<=35 vs >35).

The main discordant result in the new analysis was that a nominally significant treatment effect on disease progression was observed in high-activity/interferon-resistant subjects receiving DAC HYP, compared to those receiving IFN β-1a. Recall that, in the original presentation of the progression results in this study by the prospectively specified analysis method, there was no significant difference between the treatment groups: the hazard ratio for confirmed progression was 0.84 (DAC HYP/IFN β-1a), but the 95% CI included the possibility that progression was increased with DAC HYP (95% CI: 0.66 to 1.07). Furthermore, the hazard ratios in subgroups defined on the basis of high or low disease activity were very similar (0.80 and 0.83), as shown in the table excerpt below.

Table 24. Table excerpt of 3-month disability progression (increase in EDSS) by baseline disease activity, Study 205MS301

	IFN beta-1a 30 mcg		DAC HYP 150 mg			
	n	Estimated proportion of subjects with progression at 144 weeks (a)	n	Estimated proportion of subjects with progression at 144 weeks (a)	Hazard ratio and 95% CI (b)	P-value (b)
Disease activity at baseline						
High	204	0.201	184	0.151	0.80 (0.47, 1.38)	0.4303
Low	713	0.203	723	0.161	0.83 (0.63, 1.09)	0.1780

The post hoc shift of interferon-resistant subjects into the high-disease-activity subgroup improved the hazard ratio in this subgroup but worsened the hazard ratio for the low-disease-activity subgroup. This is not a particularly surprising result. To some extent, the hazard ratio was improved because a higher proportion of high-activity IFN β-1a subjects progressed when the subgroup was redefined to include a history of interferon resistance (24.0% progressed instead of 20.1%).

This discordant result has a weaker claim to validity than the original subgroup analysis, because it is post-hoc. After suitable adjustments for multiplicity, and with adequate appreciation of the post hoc nature of this analysis, these discordant results cannot be considered statistically robust despite nominally significant estimates for the unadjusted p-values. Furthermore, even if the finding were accepted as significant, this analysis relies on a specific definition of high activity that includes a poor response to interferon, and therefore merely suggests that DAC HYP may be more useful than beta interferon in subjects with demonstrated resistance to beta interferon. At best, this would be a weak result, and would represent a biased approach to a head-to-head comparison, focussing on subjects known to be resistant to the active comparator.

This supplementary analysis does not substantially alter the efficacy conclusions. The benefit of DAC HYP over once-weekly IFN β-1a is primarily evident in relapse rate and radiological endpoints. The evidence that DAC HYP reduces relapses in both high- and low-activity subgroups is robust, and suggests a clinically worthwhile effect. The endpoint of progression showed a trend in favour of DAC HYP in both studies, and it is possible to strengthen this finding with a number of post hoc approaches, including the enrichment of one subgroup with subjects known to be resistant to the active comparator. The placebo-controlled data failed to show a statistically significant benefit on progression, but a nominally significant p-value was obtained (if issues of multiplicity of dose groups are ignored). The cumulative effect of multiple marginal results suggests that DAC HYP is likely to have a clinically worthwhile benefit on progression, but this is yet to be confirmed in a rigorous manner. This new post hoc analysis does not substantially alter that conclusion, but merely suggests that efficacy relative to IFN β-1a may be better if subjects have already demonstrated that they respond poorly to IFN β-1a. If confirmed in a suitable prospective study, this would in turn suggest that DAC HYP could be useful as second-line or rescue therapy, which is likely to be the therapeutic role it occupies anyway.

7.2.3. EMA question 95

'The statistical and clinical significance of patient reported outcomes (PRO) and INEC confirmed disability progression are not robust. Please compare to current available treatments, considering both patients with high versus low disease activity.'

This is an important question, asking the sponsor to place the progression results in the broader context of current available treatments, but it should be noted that the subdivision of patients into 'high versus low disease activity' is not a simple exercise. As discussed in the discussion of EMA Question 94 above, one of the definitions of high disease activity included in the supplementary data was based on beta interferon treatment failure (implying interferon resistance) in addition to other measures of disease activity.

7.2.3.1. Sponsor's response

The sponsor begins their response with a statement that does not appear fully supported by the evidence:

> *'All measures of disease progression assessed in the DAC HYP clinical trials provide clinically meaningful evidence that DAC HYP-treated subjects experience less disease progression than subjects treated with placebo or IFN β-1a. Although some of the trends do not reach statistical significance, the consistency of the results provides reassurance that they are not due to chance.'*

A reduction in disease progression can only be 'clinically meaningful' if it actually exists, and in Study 205MS301 the 95% CI for the difference in disease progression was consistent with their being zero difference between DAC HYP and IFN β-1a, or even a slight inferiority of DAC HYP versus IFN β-1a. It is not the case that a significant benefit was shown but a few trends failed to reach significance, as implied in the sentences quoted above: no significant benefit was demonstrated, but a few analysis methods create nominally significant p-values if issues of multiplicity are ignored.

The sponsor's appeal to 'consistency of the results' has partial merit. A series of independent trends can sometimes, in aggregate, provide strong clues of a non-random effect, even though they are individually not significant. Furthermore, there was a broad consistency between the pivotal studies, with DAC HYP showing a strong trend to superiority over placebo at both doses (with nominally significant p-values at one dose) and a weaker trend over IFN β-1a.

Despite these consistencies across studies, the sponsor's expression 'reassurance that they are not due to chance' overstates the strength of the evidence. The question of whether chance alone could give rise to an observation can be partially quantified, and indeed this is the whole point of statistical hypothesis testing.

The sponsor's response then mentions the MSIS-29 scale, and proceeds to provide evidence of benefit with DAC HYP on MSIS-29 scores. The MSIS-29 scale is not a measure of disability progression and is not directly relevant to the question being addressed but the sponsor argues that it correlates with disease progression:

> *'The endpoints of the MSIS-29 physical scale (as a patient-reported outcome (PRO)) and disability progression by EDSS were assessed by subjects' baseline level of disease activity and support the overall conclusion that DAC HYP 150 mg offers a consistent advantage over both placebo and active comparators… The MSIS-29 physical scale has been shown to be a valid and sensitive PRO measure, which correlates highly with the EDSS (r = 0.704) and the multiple sclerosis functional composite (MSFC) (r = 0.577).'*

The link between MSIS-29 scores and disability progression is not sufficiently tight that one measure can be used as a surrogate for another and the fact that the two measures have been shown to be correlated in subjects not on DAC HYP does not allow robust inferences to be made about whether DAC HYP might differentially modify each measure.

The MSIS-29 results in Study 205MS301 have already been discussed in the description of that study, where it was noted that a significant benefit was achieved relative to IFN β-1a, but it was unclear to what extent this reflected tolerability issues with IFN β-1a. The sponsor's new comments on the MSIS-29 scores in their response to EMA Question 95 are reproduced below. Overall, there was a benefit demonstrated for this measure, and it appears plausible that this was, in part, due to benefits in disease progression. If the results for disease progression had been positive in Study 205MS301, the MSIS-29 results would have given the positive progression results some extra plausibility; they do not, however, overturn the fact that negative results were obtained in the direct analysis of disease progression.

> *'In Study 205MS201, DAC HYP 150 mg resulted in a significant decrease in the proportion of subjects with a clinically meaningful worsening (defined as an increase of ≥ 7.5 points) on the*

MSIS-29 physical scale compared to placebo (odds ratio = 0.56 (95% CI: 0.35 to 0.88; p = 0.0125]). Among the subjects with high disease activity at baseline in Study 201, the proportion of subjects who experienced a clinically meaningful decline on the MSIS-29 physical scale was higher in the placebo group (38%) than in the DAC HYP 150 mg group (22%, p = 0.1104). Amongst subjects with low disease activity at baseline, a similar benefit was seen for the DAC HYP 150 mg arm: 30% of subjects in the placebo group experienced worsening compared with 20% of subjects in the DAC HYP 150 mg group (p = 0.0462) (Table 25, below).

Similarly, in Study 205MS301, DAC HYP 150 mg prevented clinically meaningful (≥ 7.5 point increase) decline on the MSIS-29 physical score, a predefined secondary study endpoint, compared with IFN β-1a (odds ratio = 0.76 (95% CI: 0.60 to 0.95; p = 0.0176)). Among subjects with high disease activity at baseline, a greater proportion of subjects in the IFN β-1a group experienced worsening as compared with the DAC HYP group (25% versus 19%, odds ratio = 0.69, p = 0.0409). The trend in subjects with low disease activity was consistent with that of the high disease activity subgroup, but the difference between the treatments was less pronounced (22% versus 19%; odds ratio = 0.80; p = 0.1534) (Table 26, below)

The remainder of the sponsor's response addressed the EMA's question more directly, by considering the progression results of the two pivotal studies alongside studies with other disease-modifying MS treatments:

'Slowing of disability progression remains one of the most important goals of MS therapy. Although numerous agents have been shown to have a clinically meaningful and statistically significant improvement over placebo in controlled trials (Tables 25 and 26, below) no approved agent has demonstrated a consistent benefit over an active comparator agent (Tables 27 and 28, below).'

This is a fair statement and it describes the current literature reasonably well. The tables provided with the response are reproduced below (an error in the first table, in which placebo and DAC HYP results were transposed, has been fixed by the evaluator). The placebo-controlled results for DAC HYP compare favourably with other active treatments in their respective placebo-controlled studies: DAC HYP at the proposed dose was associated with the most favourable hazard ratio in the table (0.43), and the nominal p-value (p = 0.0211) compares favourably with most of the other disease-modifying agents. The placebo-controlled results for '6-Month' (24-Week) confirmed progression were even stronger, with a low hazard ratio (0.24), and the nominal p-value appears to show a high degree of significance (p = 0.0037). Against this, it should be noted that the 150 mg dose group was the secondary dose group in Study 205MS201, and a hierarchical testing procedure was used to control for multiplicity. By this procedure, no significance was demonstrated for the 150 mg dose group, and the cited p-values should be considered spurious. On the other hand, if the nominal p-values for 150 mg were doubled, to approximate the effects of assessing two doses in the same study, they would remain nominally significant. The results for 24-week confirmed progression have not been corrected for multiplicity, but doubling to account for multiplicity of doses (150 mg and 300 mg) and doubling again to account for multiplicity of confirmation times (12 and 24 weeks) would produce a nominal p-value of 0.015, well below the traditional cut-off of 0.05.

The results for the active-controlled studies show that a significant benefit versus an active comparator has not often been achieved. Fingolimod failed to show a significant benefit for 3-month confirmed disability against IFN β-1a, and alemtuzumab showed a significant benefit over IFN β-1a (44mcg three times a week, not 30mcg weekly) in one study but not another. In this respect, the results for DAC HYP are no worse than for other agents against active comparators: DAC HYP failed to show a significant benefit for the main prospective endpoint of 12 week confirmed progression, but it did show a significant benefit for the supportive, tertiary endpoint of 24 week confirmed progression.

In their response to this question, the sponsor also included results based on the subgroup analysis already discussed in relation to EMA Question 94, in which high disease activity was redefined to involve, not just high disease activity, but treatment failure on beta interferon. This post hoc analysis produces nominal hazard ratios and p-values that compare even more favourably to other disease-modifying agents, but the validity of this post hoc approach is questionable for the reasons discussed earlier.

Table 25. Impact on 3-month (12-week) confirmed disability progression of MS therapies in placebo controlled trials

Agent	% progression active arm	% progression placebo	Assessment Timepoint	Hazard ratio (active: placebo)	% reduction vs. placebo	P value vs Placebo	Reference
IFN β-1b SC	20	28%	3 years		29%	0.161	[The IFNB Multiple Sclerosis Study Group 1993]
Glatiramer acetate	21.6	24.6	2 years		12%	NS	[Johnson 1995]
Mitoxantrone 12 mg/m^2	8	22	2 years		63.6%	0.036	[Hartung 2002]
Natalizumab	17	29	2 years	0.58		<0.001	[Polman 2006]
Fingolimod 0.5 mg	17.7	24.1	2 years	0.70		0.02	[Kappos 2010]
Fingolimod 0.5 mg	25.3	29	2 years	0.83		0.227	[Calabresi 2014b]
Teriflunomide 7 mg	21.7	27.3	2 years	0.76		0.08	[O'Connor 2011]
Teriflunomide 14 mg	20.2	27.3	2 years	0.70		0.03	[O'Connor 2011]
Dimethyl fumarate 240 mg BID	13	17	2 years	0.79		0.25	[Fox 2012]
Dimethyl fumarate 240 mg BID	16	27	2 years	0.62		0.005	[Gold 2012]
Pegylated IFN β-1a 125 ug q4wk	6.8	10.5	1 year	0.62		0.0380	[Calabresi 2014a]
Peginterferon β-1a 125 ug q2wk	6.8	10.5	1 year	0.62		0.0383	[Calabresi 2014a]
DAC HYP 150 mg*	5.9	13.3	1 year	0.43		0.0211	CSR 205MS201, Table 27

Note active and placebo values were switched in Sponsor's original table for DAC HYP.

Table 26. Impact on 6-month confirmed disability progression of MS therapies in placebo controlled trials

Agent	% progression active arm	% progression placebo	Assessment Timepoint	Hazard ratio (active: placebo)	% reduction active vs placebo	P value vs placebo	Reference
IFN β-1a IM	21.9	34.9	2 years		37% progression	0.02	[Jacobs 1996]
Natalizumab			2 years	0.46		<0.001	[Polman 2006]
Fingolimod 0.5 mg	17.7	24.1	2 years	0.63		0.01	[Kappos 2010]
DAC HYP 150 mg	2.6	11.1	1 year	0.24		0.0037	205MS201 CSR Table 89

Table 27. Impact on 3-month confirmed disability progression of MS therapies in active controlled trials

Agent	Comparator	% progression active arm	% progression control arm	Assessment Timepoint	Hazard ratio (active: comparator)	% reduction active vs comparator	P value	Reference
Fingolimod 0.5 mg	IFN β-1a	5.9	7.9	1 year		25.3%	0.25	[Cohen 2010]
DAC HYP 150 mg	IFN β-1a	12	14	2 years	0.84		0.1575	CSR 205MS301, Table 31

Table 28. Impact on 6-month confirmed disability progression of MS therapies in active controlled trials

Agent	Comparator	% progression active arm	% progression control arm	Assessment Timepoint	Hazard ratio (active: comparator)	% reduction active vs comparator	P value	Reference
Alemtuzumab	IFN β-1a 44 mcg tiw	13	20	2 years	0.58		0.0084	[Coles 2012]
Alemtuzumab	IFN β-1a 44 mcg tiw	8	11	2 years	0.70		0.22	[Cohen 2012]
DAC HYP 150 mg	IFN β-1a	9.2	12.1	2 years	0.73		0.0332	CSR 205MS301, Table 32

* Comparison across clinical trials is not possible due to differences in patient populations, definitions of endpoints, observation period and analytic methods used.

7.2.3.2. Evaluator's conclusion

The sponsor's response to this question confirms that it has generally been difficult to show that disease-modifying agents have a substantial benefit in reducing disease progression, relative to active controls. The evidence that DAC HYP reduced progression is inconclusive, from a statistical perspective, but the absolute magnitude of the observed trends is favourable when compared to other new agents.

7.3. Other efficacy studies

The original CER described a number of minor studies, including extension studies (Studies 205MS202, 205MS203, 205MS303). Only Study 205MS202 was blinded. These studies shared a problem with most extension studies, in that they are subject to a number of biases due to prior exposure to study medication, non-random entry into the extension cohort, and incomplete follow-up.

Overall, these minor studies were broadly consistent with the pivotal studies, but lacked the statistical power and methodological rigour required to modify conclusions drawn directly from the pivotal studies alone. The studies were broadly reassuring in that there was no obvious waning of efficacy with continued treatment. The original CER should be consulted for details.

7.4. Efficacy comparisons with other disease-modifying agents

The submitted evidence suggests that DAC HYP has superior efficacy in comparison to once-weekly IFN β-1a, and broadly similar efficacy to other new disease-modifying agents.

In the pivotal placebo-controlled study, Study 205MS201, the observed reduction in relapse rate (approximately 50 to 54%, depending on which dose group is considered) resembled the reported reductions in ARR observed in other recent pivotal placebo-controlled studies of different disease-modifying agents, including fingolimod and dimethyl fumarate. For the pivotal placebo-controlled fingolimod trial, the ARR was 0.18 in the active group, compared to 0.40 in the placebo group, a relative reduction 55% ($p < 0.001$, approved PI for fingolimod). In the pivotal study of dimethyl fumarate, the reduction in ARR was also similar: 'The annualised relapse rate at 2 years was 0.17 in the twice-daily BG-12 group and 0.19 in the thrice-daily BG-12 group, as compared with 0.36 in the placebo group, representing relative reductions of 53% and 48% with the two BG-12 regimens, respectively ($p < 0.001$ for the comparison of each BG-12 regimen with placebo).'[5]

In a direct comparison of DAC HYP with once-weekly IFN β-1a, in Study 205MS301, the adjusted annualized relapse rate in the IFN β-1a group was 0.393 relapses/year, compared to a rate of 0.216 relapses/year in the DAC HYP group (95% CIs: 0.353 to 0.438 in the IFN β-1a treatment group and 0.191 to 0.244 in the DAC HYP treatment group). These results correspond to a relative reduction of 45% in ARR ($p < 0.0001$) with DAC HYP, compared to IFN β-1a.

DAC HYP was not shown to have a significant effect on disease progression in either of the pivotal studies, but there were favourable trends in both studies and the observed hazard ratios were broadly consistent with other disease modifying agents, as discussed.

7.5. Evaluator's overall conclusions on clinical efficacy

The sponsor provided a summary table of the key results of the two pivotal studies, and this table is reproduced below. It should be noted that the p-values flagged in the table as 'nominal' should not be considered statistically significant, indeed by a strict application of the closed testing procedure, these values should not even have been calculated or reported. Also, the cited p-values do not include any correction for multiplicity. In particular, despite the nominal

p-value of 0.0211 cited for sustained disability progression in Study 205MS201, a significant benefit on progression cannot be inferred.

The table also uses relative risk estimates that have been inflated by the practice of using instantaneous hazard ratios to estimate relative risk, as discussed previously in this report (see Sections: Results of Efficacy Outcomes for both pivotal studies).

Finally, the benefits of DAC HYP have only been demonstrated in subjects with RRMS who satisfied the entry criteria for the two pivotal studies. Extrapolation to a broader population is not warranted.

With these limitations in mind, the evaluator concludes that the following efficacy benefits are supported by the evidence:

- DAC HYP at a dose of 150 mg or 300 mg SC 4-weekly reduced annualised relapse rate by 50 to 54%, relative to placebo ($p \leq 0.0002$)
- DAC HYP at the proposed dose of 150 mg SC 4-weekly reduced relapse rate by 45%, relative to once-weekly IFN β-1a ($p < 0.0001$)
- DAC HYP at the proposed dose reduced the proportion of subjects relapsing by 44 to 47%, relative to placebo, and by 34 to 35%, relative to IFN β-1a, depending on the duration of follow-up. (Note: this is less benefit than claimed by the sponsor, for calculations see Sections: Results of Efficacy Outcomes for both pivotal studies)
- DAC HYP at the proposed dose reduced the number of new Gd-enhancing lesions by 69 to 78%, relative to placebo ($p < 0.0001$)
- DAC HYP at the proposed dose reduced the number of new or newly enlarging T2 lesions by 70 to 79% relative to placebo ($p < 0.0001$), and by 54% relative to IFN β-1a ($p < 0.0001$)
- Compared to placebo, DAC HYP showed a trend to benefits in quality of life at the proposed dose, as estimated by the MSIS-29 physical impact score, but by the closed testing procedure failed to achieve significance, and trends were inconsistent across dose groups
- Compared to beta interferon, DAC HYP showed significant superiority in quality of life, as estimated by the MSIS-29 physical impact score
- DAC HYP is associated with a strong trend to reduced disability progression
- DAC HYP produced a broadly similar benefit across all major subgroups in the study population
- DAC HYP at the proposed dose has better efficacy, relative to IFN β-1a, in a population enriched for subjects with proven resistance to IFN β-1a, and in this population a nominally significant post hoc p-value can be obtained for the endpoint of progression
- DAC HYP at the proposed dose has a broadly similar efficacy to other new disease-modifying agents
- DAC HYP has not been studied in subjects with overt SPMS, and its efficacy in this population is unknown

Despite the fact that the supplementary evaluator and the sponsor have drawn different conclusions about the statistical robustness of the progression data, these efficacy results are considered satisfactory. The supplementary evaluator does not believe that a clear benefit on progression endpoints should be an absolute requirement for a new disease-modifying agent in MS. In subjects with RRMS, a large proportion of disability progression is due to damage sustained during relapses, and preventing relapses is a worthwhile achievement in its own right, provided that there is at least no adverse effect on progression. Although the data do not provide robust confirmation of a benefit for progression endpoints, there is a consistency across

multiple different analyses that, in aggregate, strongly suggest that DAC HYP has a favourable effect on progression, and at least DAC HYP appears highly unlikely to have an adverse effect on progression. Coupled with strong evidence of a reduced relapse rate, this is sufficient to support the claim of efficacy in RRMS.

The efficacy of DAC HYP in subjects with SPMS has not been characterised, and there is currently no basis for recommending this treatment in subjects with SPMS.

The lowest dose of DAC HYP capable of producing a substantial reduction in relapse rate has not been established.

Table 29. Primary, secondary, and selected tertiary efficacy endpoints in the DAC HYP pivotal studies 205MS201 and 205MS301, DAC HYP 150 mg

Primary, Secondary, and Selected Tertiary Efficacy Endpoints in the DAC HYP Pivotal Studies 205MS201 and 205MS301 - DAC 150mg

Efficacy Endpoints	Study 205MS201 reduction vs. Placebo	Study 205MS301 reduction vs. IFN ß-1a
Clinical		
Annualized relapse rate*	53.9% p<0.0001	45.0% p<0.0001
Proportion of subjects relapsed	54.9% p<0.0001**	40.9% p<0.0001*#
Disability progression sustained for 12 weeks as measured by EDSS Score	56.6% p=0.0211***	16.1% p=0.1575**
Disability progression sustained for 24 weeks as measured by EDSS Score	76.4% p=0.0037	27.0% p=0.0332***
The change from baseline in MSIS-29 physical score	-4.27 p=0.0008*#	-2.09 p=0.0008***
Proportion of subjects with a >=7.5 point worsening from baseline in the MSIS-29 physical impact score	44.2% p=0.0125	24.2% p=0.0176*#
MRI endpoints		
The number of new Gd-enhancing lesions between weeks 8 and 24	69.5% p<0.0001**	n/a
The number of new or newly enlarging T2 hyperintense lesions	70.2% p<0.0001**	54.4% p<0.0001**
The number of Gd-enhancing lesions at week 96	n/a	75.0% p<0.0001***
The number of new T1 hypointense lesions at week 96	n/a	51.8% p<0.0001***

* Primary endpoint; Secondary endpoints for each study were ranked and a sequential closed testing procedure was employed to control for type I error at 5%; Tertiary endpoints were pre-specified in each statistical analysis plan but the analyses were not adjusted for multiplicity.
** Secondary endpoint
*# Secondary endpoints that was not tested based on the closed testing procedure. Nominal p-values are presented.
*** Tertiary endpoint

Note: the evaluator does not agree with all of the figures of this table – see text for details)

8. Clinical safety

8.1. Summary of clinical safety from the CER

The First Round Clinical Evaluation Report (CER) summarised the safety of DAC HYP as follows:

- AEs were common in subjects treated with DAC HYP. Treatment-related adverse events (TRAE) occurred in about 22% of subjects treated with DAC HYP, the most common being injection site pain, influenza-like illness, headache, alanine transaminase (ALT) increased, aspartate transaminase (AST) increased, liver function tests (LFT) abnormal, gamma-glutamyl transaminase (GGT) increased, nasopharyngitis, pyrexia, injection site erythema, injection site bruising, upper respiratory tract infection, pharyngitis, MS relapse, fatigue, rash, eczema, nausea, lymphadenopathy, and lymphopaenia. The majority were mild to moderate and were manageable with standard treatment or interruption or discontinuation of DAC HYP.
- There were two deaths attributed to DAC HYP. One was a case of autoimmune hepatitis following planned washout and re-initiation of DAC HYP. The second was a case of bacteraemia, following an exfoliative rash leading to the development of a psoas abscess, emboli and bowel ischaemia.
- The case of hepatitis lead to more intensive monitoring in the clinical study programs. Though derangements in LFTs remained common in the DAC HYP studies, these cases were managed by interruption or discontinuation of treatment. No further 'Hy's Law' cases were seen, and there were no further episodes of hepatitis. It is considered that the risk can be adequately managed in the post-market environment with a program that frequently monitors LFTs and the provision of adequate advice with regard to managing derangements.
- The second death was related to DAC HYP but appears to have been as a secondary consequence of the adverse event of skin rash. It is unclear whether this case was true Steven's-Johnson syndrome (SJS) but, at the very least, it was a case of a severe skin hypersensitivity reaction.
- Skin reactions were common treatment emergent adverse events (TRAE) and occurred in about 37% of subjects in the active control study; 2% of cutaneous adverse events met the criteria for serious. Typically, the cutaneous adverse events were mild to moderate in nature and resolved with treatment. The serious cases were treated with systemic corticosteroids; this should be reflected in the PI.
- A low incidence of colitis was seen in patients treated with DAC HYP. The colitis largely resolved after DAC HYP was discontinued. The mechanism, optimal treatment and long-term management remain unknown.
- An excess of mild to moderate depression in subjects treated with DAC HYP was seen in the placebo-controlled study. The incidence of depression appears to be no worse than that for IFN β-1a which is a standard treatment for MS. DAC HYP should be contraindicated in patients with a recent history of severe depression.

9. Supplementary safety data

As noted in Section 7.2.2, EMA Question 94 (p60) the sponsor reanalysed Treatment Emergent Adverse Events (TEAEs) according to their new, post hoc definition of high disease activity. In this analysis, the number of subjects with an AE was broadly similar in subjects with high disease activity and low disease activity, as shown in the tables below. In Study MS201, there was a slight excess of AEs in subjects with high disease activity, but there was no consistent difference across treatment groups. For subjects with serious AEs, the placebo group had a higher incidence than either of the active dose groups (32% and 25% in high activity and low activity subgroups, respectively). There was no consistent pattern amongst active groups, with the 150 mg dose group showing more SAEs in the high activity subgroup (19% and 14%), and

the 300 mg dose group showing a slight excess of SAEs in the low/unknown activity subgroup (15% and 18%).

In Study MS301, the proportion of subjects with a TEAE was similar across both treatment arms and in both high and low activity subgroups. For SAEs, there was a similar incidence in each treatment arm, with the exactly the same incidence in the high activity subgroups (28% for IFN β-1a and for DAC HYP), but a slightly lower incidence in the low activity subgroups (16% for IFN β-1a and 22% for DAC HYP).

For hepatic events and for abnormal LFTs, there was no overall pattern in relation to disease activity (as shown in the tables below). Although there is evidence of some hepatic risk with DAC HYP treatment, discussed in more detail in later sections, this risk does not appear to be particularly prominent in subgroups defined on the basis of disease activity.

Similarly, for cutaneous events, there was a clear excess in recipients of DAC HYP, relative to placebo or IFN β-1a, but the problem was not prominent in the high or low disease activity subgroups.

9.1. Study MS201

Table 30. Summary of TEAEs by Disease Activity Subgroup, Study MS201

	Placebo		150 mg DAC HYP		300 mg DAC HYP	
	High	Low/Unknown	High	Low/Unknown	High	Low/Unknown
Number of subjects in the 205MS201 safety population	37 (100)	167 (100)	48 (100)	160 (100)	39 (100)	170 (100)
Number of subjects with an event	31 (84)	130 (78)	41 (85)	110 (69)	29 (74)	130 (76)
Number of subjects with a serious event	12 (32)	41 (25)	9 (19)	23 (14)	6 (15)	30 (18)
Number of subjects discontinuing treatment due to an event	0	2 (1)	1 (2)	6 (4)	2 (5)	7 (4)
Number of subjects withdrawing from study due to an event	0	1 (<1)	1 (2)	3 (2)	0	3 (2)
Number of subjects with an infection	18 (49)	72 (43)	27 (56)	77 (48)	19 (49)	93 (55)
Number of subjects with a serious infection	0	0	2 (4)	4 (3)	0	3 (2)
Number of subjects with a cutaneous event	3 (8)	24 (14)	7 (15)	31 (19)	7 (18)	38 (22)
Number of subjects with a serious cutaneous event	0	0	1 (2)	1 (<1)	1 (3)	2 (1)
Number of subjects with a hepatic event	2 (5)	10 (6)	5 (10)	14 (9)	2 (5)	16 (9)
Number of subjects with a serious hepatic event	0	0	0	2 (1)	1 (3)	0
Number of subjects with post-baseline ALT or AST > 5xULN	0	1 (<1)	2 (4)	7 (4)	1 (3)	7 (4)

Number of subjects in the 205MS201 safety population	37 (100)	167 (100)	87 (100)	330 (100)
Number of subjects with an event	31 (84)	130 (78)	70 (80)	240 (73)
Number of subjects with a serious event	12 (32)	41 (25)	15 (17)	53 (16)
Number of subjects discontinuing treatment due to an event	0	2 (1)	3 (3)	13 (4)
Number of subjects withdrawing from study due to an event	0	1 (<1)	1 (1)	6 (2)
Number of subjects with an infection	18 (49)	72 (43)	46 (53)	170 (52)
Number of subjects with a serious infection	0	0	2 (2)	7 (2)

9.2. Study MS301

Table 31. Summary of TEAEs by Disease Activity Subgroup, Study MS301

	IFN beta-1a 30 mcg		DAC HYP 150 mg	
	High	Low/Unknown	High	Low/Unknown
Number of subjects in the 205MS301 safety population	386 (100)	536 (100)	358 (100)	561 (100)
Number of subjects with an event	359 (93)	483 (90)	333 (93)	505 (90)
Number of subjects with a serious event	107 (28)	87 (16)	99 (28)	122 (22)
Number of subjects discontinuing treatment due to an event	65 (17)	47 (9)	50 (14)	92 (16)
Number of subjects withdrawing from study due to an event	35 (9)	31 (6)	23 (6)	41 (7)
Number of subjects with an infection	229 (59)	294 (55)	249 (70)	346 (62)
Number of subjects with a serious infection	8 (2)	7 (1)	20 (6)	20 (4)
Number of subjects with a cutaneous event	78 (20)	98 (18)	148 (41)	196 (35)
Number of subjects with a serious cutaneous event	1 (<1)	0	3 (<1)	11 (2)
Number of subjects with a hepatic event	63 (16)	67 (13)	53 (15)	91 (16)
Number of subjects with a serious hepatic event	3 (<1)	1 (<1)	2 (<1)	4 (<1)
Number of subjects with post-baseline ALT or AST > 5xULN	16 (4)	15 (3)	16 (4)	43 (8)

NOTE 1: Numbers in parentheses are percentages.
2: Includes events started between the First Dosing Date and up to 180 days after Last Dosing Date.
3: High disease activity at baseline is defined as 1) patients with >= 2 relapses in the previous year and >= 1 Gd lesion at baseline, or 2) patients who have failed to respond to at least one year of treatment of beta-interferon, having had >= 1 relapse in the previous year while on therapy, and >= 9 T2 lesions or >=1 Gd lesion, or patients having an unchanged or increased relapse rate in the prior year as compared to the previous 2 years. Otherwise disease activity is 'Low/Unknown'.

Table 32. Maximum Liver Function Test Values by Disease Activity Subgroup, Study MS301

Laboratory Parameters	Criterion	IFN beta-1a 30 mcg		DAC HYP 150 mg	
		High	Low/Unknown	High	Low/Unknown
ALT	Total n	380 (100)	534 (100)	358 (100)	559 (100)
	<=1 xULN	208 (55)	265 (50)	226 (63)	314 (56)
	>1 xULN	172 (45)	269 (50)	132 (37)	245 (44)
	>=3 xULN	37 (10)	39 (7)	30 (8)	57 (10)
	>5 xULN	15 (4)	15 (3)	15 (4)	38 (7)
	>10 xULN	6 (2)	5 (<1)	8 (2)	16 (3)
	>20 xULN	3 (<1)	1 (<1)	4 (1)	4 (<1)
AST	Total n	380 (100)	534 (100)	358 (100)	559 (100)
	<=1 xULN	256 (67)	349 (65)	253 (71)	363 (65)
	>1 xULN	124 (33)	185 (35)	105 (29)	196 (35)
	>=3 xULN	16 (4)	18 (3)	19 (5)	44 (8)
	>5 xULN	7 (2)	7 (1)	11 (3)	26 (5)
	>10 xULN	2 (<1)	3 (<1)	7 (2)	7 (1)
	>20 xULN	1 (<1)	1 (<1)	2 (<1)	4 (<1)
Worst of ALT or AST	Total n	380 (100)	534 (100)	358 (100)	559 (100)
	<=1 xULN	190 (50)	246 (46)	207 (58)	287 (51)
	>1 xULN	190 (50)	288 (54)	151 (42)	272 (49)
	>=3 xULN	40 (11)	40 (7)	32 (9)	64 (11)
	>5 xULN	16 (4)	15 (3)	16 (4)	43 (8)
	>10 xULN	6 (2)	6 (1)	9 (3)	16 (3)
	>20 xULN	3 (<1)	1 (<1)	4 (1)	4 (<1)
Total Bilirubin	Total n	380 (100)	534 (100)	358 (100)	559 (100)
	<=1 xULN	350 (92)	495 (93)	322 (90)	491 (88)
	>1 xULN	30 (8)	39 (7)	36 (10)	68 (12)
	>1.5 xULN	13 (3)	5 (<1)	15 (4)	21 (4)
	>2 xULN	2 (<1)	2 (<1)	7 (2)	12 (2)
	>3 xULN	0	0	1 (<1)	2 (<1)
	>10 xULN	0	0	0	0
GGT	Total n	380 (100)	534 (100)	358 (100)	559 (100)
	<=1 xULN	317 (83)	432 (81)	294 (82)	445 (80)
	>1 xULN	63 (17)	102 (19)	64 (18)	114 (20)
	>2.5 xULN	14 (4)	23 (4)	13 (4)	31 (6)
	>5 xULN	1 (<1)	5 (<1)	3 (<1)	11 (2)
	>20 xULN	0	0	1 (<1)	0
Alkaline Phosphatase	Total n	380 (100)	534 (100)	358 (100)	559 (100)
	<=1 xULN	360 (95)	512 (96)	323 (90)	501 (90)
	>1 xULN	20 (5)	22 (4)	35 (10)	58 (10)
	>1.5 xULN	3 (<1)	5 (<1)	8 (2)	17 (3)
	>2.5 xULN	0	0	3 (<1)	3 (<1)
	>5 xULN	0	0	1 (<1)	0
	>20 xULN	0	0	0	0
Elevations in ALT or AST >=3xULN accompanied by concurrently elevated total bilirubin	Total n	380 (100)	534 (100)	358 (100)	559 (100)
	>1.5 xULN	2 (<1)	1 (<1)	2 (<1)	8 (1)
	>2 xULN	0	1 (<1)	2 (<1)	5 (<1)

NOTE 1: Numbers in parentheses are percentages.
2: Includes data up to 180 days after Last Dosing Date.
3: High disease activity at baseline is defined as 1) patients with >= 2 relapses in the previous year and >= 1 Gd lesion at baseline, or 2) patients who have failed to respond to at least one year of treatment of beta-interferon, having had >= 1 relapse in the previous year while on therapy, and >= 9 T2 lesions or >=1 Gd lesion, or patients having an unchanged or increased relapse rate in the prior year as compared to the previous 2 years. Otherwise disease activity is 'Low/Unknown'.
Abbreviations: ULN=upper limit of normal.

9.3. Safety issues flagged in the US product information

Zinbryta was approved in the US during preparation of this report. Because of concerns about hepatic injury and immune-mediated disorders, the US Product Information carries a black box warning, as shown below. The safety issues were considered sufficiently serious that the drug is currently only available in the US through a restricted distribution program. (Section numbers within the boxes refer to the US PI.)

9.3.1. Summary on first page

WARNING: HEPATIC INJURY INCLUDING AUTOIMMUNE HEPATITIS and OTHER IMMUNE-MEDIATED DISORDERS

See full prescribing information for complete boxed warning.

Hepatic Injury Including Autoimmune Hepatitis

ZINBRYTA can cause severe liver injury including life-threatening events, liver failure, and autoimmune hepatitis. Obtain transaminase and bilirubin levels before initiation of ZINBRYTA. Monitor and evaluate transaminase and bilirubin levels monthly and up to 6 months after the last dose (2.3, 2.4, 5.1).

ZINBRYTA is contraindicated in patients with pre-existing hepatic disease or hepatic impairment (4, 5.1).

Other Immune-Mediated Disorders

Immune-mediated disorders including skin reactions, lymphadenopathy, non-infectious colitis, and other immune-mediated disorders can occur with ZINBRYTA (5.2).

These conditions may require treatment with systemic corticosteroids or immunosuppressive medication (5.1, 5.2).

ZINBRYTA is available only through a restricted distribution program called the ZINBRYTA REMS Program (5.3).

9.3.2. Complete boxed warning

WARNING: HEPATIC INJURY INCLUDING AUTOIMMUNE HEPATITIS and OTHER IMMUNE-MEDIATED DISORDERS

Hepatic Injury Including Autoimmune Hepatitis

ZINBRYTA can cause severe liver injury including life-threatening events, liver failure, and autoimmune hepatitis. In clinical trials, 1 patient died due to autoimmune hepatitis. Liver injury, including autoimmune hepatitis, can occur at any time during treatment with ZINBRYTA, with cases reported up to 4 months after the last dose of ZINBRYTA.

ZINBRYTA is contraindicated in patients with pre-existing hepatic disease or hepatic impairment *[see Contraindications (4) and Warnings and Precautions (5.1)]*.

Prior to starting ZINBRYTA, obtain serum transaminases (ALT and AST) and bilirubin levels *[see Dosage and Administration (2.3)]*.

Test transaminase levels and total bilirubin monthly and assess before the next dose of ZINBRYTA. Follow transaminase levels and total bilirubin monthly for 6 months after the last dose of ZINBRYTA. In case of elevation in transaminases or total bilirubin, treatment interruption or discontinuation may be required *[see Dosage and Administration (2.4) and Warnings and Precautions (5.1)]*.

Other Immune-Mediated Disorders

In addition to autoimmune hepatitis, immune-mediated disorders such as skin reactions, lymphadenopathy, and non-infectious colitis can occur in patients treated with ZINBRYTA. Overall, serious immune-mediated conditions were observed in 5% of patients treated with ZINBRYTA *[see Warnings and Precautions (5.2)].*

If a patient develops a serious immune-mediated disorder, consider stopping ZINBRYTA and refer the patient to a specialist to ensure comprehensive diagnostic evaluation and appropriate

treatment.

Some patients required systemic corticosteroids or other immunosuppressant treatment for autoimmune hepatitis or other immune-mediated disorders and continued this treatment after the last dose of ZINBRYTA [see Warnings and Precautions (5.1, 5.2)].

Because of the risks of hepatic injury, including autoimmune hepatitis, and other immune-mediated disorders, ZINBRYTA is available only through a restricted program under a Risk Evaluation and Mitigation Strategy (REMS) called the ZINBRYTA REMS Program [see Warnings and Precautions (5.3)].

9.4. Patient exposure

Patient exposure was not evaluated in the SCER.

Please also see the extract of the CER (Attachment 2).

9.5. Post-marketing experience

The sponsor's comments on post-marketing data for DAC HYP begin with the following observation: 'DAC HYP is an investigational product and has not been approved or marketed in any countries.'

The sponsor then discusses the fact that similar agents have been marketed previously, including two products containing alternative preparations of daclizumab.:

'In addition to DAC HYP, there are several other drugs that target CD25, including 2 other forms of daclizumab (Zenapax, also referred as DAC Nutley; and DAC Penzberg) and basiliximab (Simulect). Simulect and Zenapax are approved products and are indicated for prophylaxis of acute organ rejection in patients receiving renal transplant. DAC Penzberg was used only as an investigational product and has not been approved or marketed'.

Simulect has been assessed in 4 randomised, double-blind, placebo-controlled clinical studies for the prevention of renal allograft rejection.[10,11, 12, 13, 14]

It is difficult to draw any firm conclusions about the safety of DAC HYP from this indirect evidence, partly because the patient population had substantial comorbidities and they were often treated with other immunosuppressive agents. The sponsor notes that Simulect was associated with hypersensitivity reactions:

'Severe acute hypersensitivity reactions (onset within 24 hours), including anaphylaxis, have been observed on initial exposure to Simulect and/or following re-exposure after several months.

[10]Kahan B. Rajagopalan P.R., Hall M. Reduction of the occurrence of acute cellular Rejection among renal allograft recipients treated with basiliximab, a chimeric anti-interleukin-2receptor monoclonal antibody. Transplantation Vol. 67, 276–284, No. 2, January 27, 1999.

[11] Lawen JG, Davies EA, Mourad G, Oppenheimer F, Molina MG, Rostaing L, Wilkinson AH, Mulloy LL, Bourbigot BJ, Prestele H, Korn A, Girault D, on behalf of the SIMULECT International Study Group. Randomized double-blind study of immunoprophylaxis with basiliximab, a chimeric anti-interleukin-2 Receptor monoclonal antibody, in combination with Mycophenolate mofetil-containing triple therapy in Renal transplantation. Transplantation Vol. 75, 37–43, No. 1, January 15, 2003

[12]Nashan B, Moore R, Amlot P, Schmidt A-G, Abeywickrama K, Soulillou J-P, for the CHIB 201 International Study Group. Randomised trial of basiliximab versus placebo for control of acute cellular rejection in renal allograft recipients. The Lancet Vol 350 October 25, 1997

[13]Ponticelli C, Yussim A, Cambi V, Legendre C, Rizz0 G, Salvadori M, Kahn D, Kashi H, Salmela K, Fricke L, Heemann U, Garcia-Martinez J, Leohler R, Prestele H, Girault D, on behalf of the SIMULECT® Phase IV Study Group. A randomized, double-blind trial of basiliximab immunoprophylaxis plus triple therapy in kidney transplant recipients. Transplantation Vol. 72, 1261-1267, No. 7, October 15, 2001

[14]Simulect USPI 2003

Capillary leak syndrome and cytokine release syndrome have been reported during the post-marketing experience with Simulect. Simulect should be administered only in facilities equipped and staffed with adequate laboratory and supportive medical resources.'

In the absence of clear evidence that DAC HYP is less likely to cause hypersensitivity reactions, similar caution should be used with DAC HYP.

The proposed Australian PI only contains one reference to hypersensitivity: 'Zinbryta is contraindicated in patients with a history of severe hypersensitivity (for example anaphylaxis or anaphylactoid reactions) to daclizumab, or any of the excipients.' In view of the experience with Simulect, this issue should be given greater prominence. The US PI includes several references to hypersensitivity, along with the following comment: 'Zinbryta can cause anaphylaxis, angioedema, and urticaria after the first dose or at any time during treatment. Discontinue and do not re-start Zinbryta if anaphylaxis or other allergic reactions occur.'

The other forms of daclizumab for which there is trial experience include DAC Penzberg and DAC Nutley. Several studies have been performed for each of these agents, as summarised by the sponsor in their SCS. Daclizumab was associated with a greater incidence of infections and hepatobiliary dsorders than placebo, when used in the renal transplant setting and in subjects with MS. A full evaluation of this material is beyond the scope of this report, but no substantial new issues appear to arise from a review of the sponsor's brief summary. As the sponsor notes, these older daclizumab products differ from DAC HYP in terms of the cell line used to produce the antibodies, the manufacturing process and the degree of glycosolation. DAC HYP also differs from these other daclizumab agents in that it has less antibody-dependent cellular-cytotoxicity activity in functional assays.

9.6. Safety issues the potential for major regulatory impact

9.6.1. Liver toxicity

The pivotal studies suggested that DAC HYP may be associated with liver injury in a small proportion of subjects and occasionally this may be severe. Two deaths in the study program were considered potentially attributable to DAC HYP, and one of these was caused by autoimmune hepatitis (the other was caused by sepsis and ischaemic colitis occurring as a complication of skin lesions).

The sponsor provided a narrative for the hepatic death, which occurred in a woman who was exposed to DAC HYP 300 mg in two different studies (Study 205MS201 and the extension, Study 205MS202).

DAC HYP 300 mg/washout/300 mg

Subject [information redacted] (Autoimmune hepatitis; Study 205MS202): The subject was a [information redacted] female with a medical history of MS, chronic pyelonephritis, and photoallergy and no other risk factors for liver disease, who died of autoimmune hepatitis, liver failure and multiple organ failure on Day 692 (Day 315 of the extension). The subject received 13 doses of DAC HYP in Study 205MS201 and 5 doses of placebo and 4 doses of DAC HYP in Study 205MS202 with the last dose on Day 602, approximately 3 months prior to her death. The investigator assessed the death as unrelated to study treatment. The sponsor assessed the death as related because a contributory role for DAC HYP could not be excluded.

This narrative is too brief to allow a complete assessment of the likelihood that DAC HYP contributed to the death. The fact that the subject had nearly two years of exposure before dying of hepatitis and the delay between the last dose and the death, argues against a direct toxic effect of DAC HYP on the liver, but is nonetheless consistent with a potential autoimmune mechanism triggered by DAC HYP exposure.

Minor shifts in LFTs did not appear to be more common in recipients of DAC HYP. In the placebo-controlled pivotal study, Study 205MS201, shifts from normal to high in LFTs occurred in similar percentages in the placebo and DAC HYP treatment groups for ALT (29% placebo versus 28% DAC HYP), total bilirubin (8% placebo versus 7% DAC HYP), GGT (10% for both groups), and alkaline phosphatase (ALP) (4% placebo, 5% DAC HYP). There were more shifts from normal to high for AST in the DAC HYP group (19%) than in the placebo group (11%). In the active-controlled study, 205MS301, shifts in LFTs from normal to high were also similar between the IFN β-1a and DAC HYP treatment groups, including ALT (45% and 38%), AST (33% and 31%), total bilirubin (5% and 8%), GGT (16% and 17%), and ALP (4% and 9%), respectively.

In Study 205MS201, severe shifts were slightly more common in recipients of DAC HYP, but only when levels > 5 x the upper limit of normal (ULN) were considered. Although the number of patients involved was small, and the excess in the DAC HYP group was partially disguised by the presentation of the incidence data in separate bins for different levels of elevation, abnormalities were clearly more common in the DAC HYP group. The total number of subjects with worst AST or ALT > 3 x ULN was 7 out of 204 subjects (3.4%) for placebo and 30 out of 414 subjects (7.2%) for DAC HYP. This difference in distribution of outcomes approaches (p = 0.07) or achieves (p = 0.04) statistical significance by Fisher's exact test, depending on whether a two- or one-tailed test is used (without considering issues of multiplicity).

Table 33. Incidence of maximum post-baseline LFT abnormalities, Study 205MS201

Incidence of maximum post baseline liver function abnormalities in Study 201

Study 201 Parameter/ Criterion	Placebo (n=204)	DAC HYP 150 mg (n=208)	DAC HYP 300 mg (n=209)
Worst ALT or AST, n (%)			
n	204 (100)	206 (100)	208 (100)
<=1xULN	133 (65)	136 (66)	132 (63)
>1-3xULN	64 (31)	54 (26)	62 (30)
>3-5xULN	6 (3)	7 (3)	6 (3)
>5-20xULN	1 (<1)	6 (3)	6 (3)
>20xULN	0	3 (1)	2 (<1)
ALT/AST >= 3×ULN by concurrently elevated bilirubin defined as, n (%)			
n	204 (100)	206 (100)	208 (100)
>=2xULN	1 (<1)	1 (<1)	1 (<1)

NOTE: Numbers in parentheses are percentages

In Study 205MS301, an excess in moderately and severely elevated LFTs was also observed, as shown in Table 35 below. For ALT or AST ≥ 3 x ULN, the difference in incidence between IFN β-1a and DAC HYP was minimal (9% with IFN β-1a versus 10% with DAC HYP). For ALT or AST > 5 x ULN, there was a two-fold excess in the DAC HYP group (3% versus 6%), but this was based on very low patient numbers.

Table 34. Incidence of maximum post-baseline LFT abnormalities, Study 205MS301

Laboratory Parameters	Criterion	IFN beta-1a 30 mcg	DAC HYP 150 mg
Worst ALT or AST			
	Total n	914 (100)	917 (100)
	<=1 xULN	436 (48)	494 (54)
	>1 xULN	478 (52)	423 (46)
	>=3 xULN	80 (9)	96 (10)
	>5 xULN	31 (3)	59 (6)
	>10 xULN	12 (1)	25 (3)
	>20 xULN	4 (<1)	8 (<1)
Total Bilirubin			
	Total n	914 (100)	917 (100)
	<=1 xULN	845 (92)	813 (89)
	>1 xULN	69 (8)	104 (11)
	>1.5 xULN	18 (2)	36 (4)
	>2 xULN	4 (<1)	19 (2)
	>3 xULN	0	3 (<1)
	>10 xULN	0	0

NOTE 1: Numbers in parentheses are percentages.
2: Includes data up to 180 days after Last Dosing Date.
(a) Total n is the number of subjects with at least one post-baseline measurement of concurrent total bilirubin and ALT/AST.
Abbreviations: ULN = upper limit of normal.

These results should be interpreted in the context of a study that involved close monitoring of LFTs, as explained by the sponsor: '*Following the occurrence of fatal autoimmune hepatitis during Study 205MS202 approximately 3 months after discontinuation of DAC HYP, all ongoing study protocols were updated to include LFT monitoring every 4 weeks during treatment if not already required, and guidelines were added for temporarily suspending dosing for ALT or AST elevations > 3 x ULN and permanently discontinuing study treatment for confirmed elevations of ALT or AST > 5 x ULN, or for elevations of ALT or AST > 3 x ULN that lasted longer than 1 week.*'

Close monitoring of this nature would be expected to increase the sensitivity for detecting abnormal LFTs while they were still mild, and hopefully reduce the incidence of progressive hepatic injury. More severe cases of liver injury might be expected if DAC HYP were used in a clinical setting with less stringent monitoring.

Subjects satisfying Hy's law (concurrent elevation of ALT/AST ≥ 3 x ULN and bilirubin ≥ 2 x ULN) were infrequent in both pivotal studies. In Study 205MS201, three subjects (1 in the placebo group and 2 in the DAC HYP-treated groups) had elevations in liver transaminases ≥ 3 x ULN in association with bilirubin > 2 x ULN. In Study 205MS301, substantially more subjects in the DAC HYP group (7 subjects) had had elevations in liver transaminases ≥ 3 x ULN and concurrent elevation in total bilirubin > 2 x ULN than in the IFN β-1a group (1 subject). A review of the cases found only two subjects (one in each treatment group) that were felt to represent cases of Hy's Law, which requires a lack of alternative explanations in addition to the observed abnormalities in LFTs. Nonetheless, the excess in the DAC HYP group, as well as the overall excess of moderately elevated transaminases in both studies, is of concern, particularly in view of the fact that IFN β-1a has also been associated with an excess of abnormal LFTs.

On balance, reviewing all of this evidence, it appears likely that DAC HYP carries a small but significant risk of inducing hepatic injury, which can occasionally be severe, and it may have contributed to one death by this mechanism.

9.6.2. Hypersensitivity reactions

Hypersensitivity reactions to DAC HYP occurred at a low incidence in the pivotal studies, but these were occasionally severe. One subject had probable anaphylaxis in response to DAC HYP 300 mg, with hypotension, as described in the original CER (please see Attachment 2). One subject exposed to IFN β-1a also had probable anaphylaxis.

Across the full spectrum of hypersensitivity reactions, there was an excess of events in the DAC HYP treatment groups in both pivotal studies. In Study 205MS201, the incidence of hypersensitivty-related AEs was 8%, 12%, and 12% for the placebo, DAC HYP 150 mg, and

DAC HYP 300 mg groups, respectively. Most of these were skin and subcutaneous disorders (6%, 11%, and 11% in the placebo, DAC HYP 150 mg, and DAC HYP 300 mg groups, respectively), including rash (3%, 6%, and 5%, respectively) and allergic dermatitis (< 1%, 2%, and 2%, respectively). Other hypersensitivity-related AEs had an incidence of ≤ 1% for each group.

In Study 205MS301, there was also a higher incidence of AEs in the DAC HYP group (25%) than in the IFN β-1a group (12%) and many of these were skin and subcutaneous tissue disorders (11% for IFN β-1a and 23% for DAC HYP). The most common hypersensitivity-related AEs in the DAC HYP group were rash, eczema, dermatitis, allergic dermatitis, urticaria, maculo-papular rash, contact dermatitis and atopic dermatitis. All other hypersensitivity-related AEs had an incidence of ≤ 1%.

Lymphadenopathy was also more common in DAC HYP recipients and is mentioned in the US black box warning. In Study 205MS201, the incidence of lymphadenopathy was broadly similar across treatment groups (1% placebo, 2% DAC HYP 150 mg, < 1% DAC HYP 300 mg). By contrast, in Study 205MS301, the incidence of lymphadenopathy events was clearly higher in the DAC HYP group (5%) than in the IFN β-1a group (< 1%). There was also an excess of DAC HYP recipients who had AEs of lymphadenitis (1 subject (< 1%) for IFN β-1a, 13 subjects (1%) for DAC HYP). 8 subjects in the DAC HYP group but no subjects in the IFN β-1a group had lymphadenopathy-related events rated as serious (lymphadenopathy, 5 subjects; lymphadenitis, 3 subjects).

One of the two deaths classified as potentially attributable to DAC HYP was related to a skin reaction to DAC HYP, although the mechanism of death was somewhat unrelated. A rash became infected, led to sepsis, and the patient subsequently developed bowel ischaemia.

9.6.3. Infections

The pivotal studies showed a small excess of infections that is likely to represent a mildly immunosuppressive effect of DAC HYP. The excess was observed in both pivotal studies, and included both mild and severe infections.

In Study 205MS201, infections were reported in 44%, 50%, and 54% of subjects in the placebo, DAC HYP 150 mg, and DAC HYP 300 mg groups, respectively. The most common infections were upper respiratory tract infections (26% placebo versus 31% DAC HYP) and viral infections (6% placebo versus 10% DAC HYP). The incidence of infections rated as 'severe' was 0% in the placebo group, compared to 1% in the DAC HYP 150 mg group, and < 1% in the DAC HYP 300 mg group. The incidence of serious infections was 0% in the placebo group and 2% (9 subjects) in the combined DAC HYP groups. The incidences of herpes viral infections (6% placebo versus 6% DAC HYP) and candida infections (0% placebo versus < 1% DAC HYP) were not substantially different across treatment groups, but infections characterised as 'opportunistic' were only seen in in the DAC HYP group as no subjects in the placebo group versus 4 subjects (< 1%) in the combined DAC HYP treatment groups experienced potential opportunistic infections; these included oral candidiasis (3 subjects) and cytomegalovirus infection (1 subject).

In Study 205MS301, there was also an excess of infections in the DAC HYP group, but there was no clear pattern and no clear excess of potentially opportunistic infections. The overall incidence of infections was 57% and 65% in the IFN β-1a and DAC HYP groups, respectively. The most common infections were upper respiratory tract infections (39% IFN β-1a versus 44% DAC HYP) and urinary tract infections (12% in both groups). The incidence of herpes viral infections (7% IFN β-1a versus 8% DAC HYP) and candida infections (2% IFN β-1a versus 2% DAC HYP) was similar in both treatment groups. Most infections were rated as 'mild' or 'moderate' in severity, but infections rated as 'severe' were more common with DAC HYP (1% IFN β-1a and 3% DAC HYP). 3 subjects (< 1%) in the IFN β-1a group and 5 subjects (< 1%) in

the DAC HYP group discontinued treatment due to infections. The incidence of serious infections was also increased with DAC HYP (2% IFN β-1a versus 4% DAC HYP).

Most of the serious infections were typical of the infections that can occur in a general population of MS subjects: serious infections that occurred in more than 3 subjects in the IFN β-1a or DAC HYP group included urinary tract infection (2 subjects IFN β-1a, 8 subjects DAC HYP), and pneumonia (2 subjects IFN β-1a, 5 subjects DAC HYP).

Unlike the placebo-controlled study, the incidence of potential opportunistic infections was similar between the 2 groups (2% in both groups). The most common potential opportunistic infections were candida infections, which were similar in incidence across the IFN β-1a (2%) and DAC HYP (2%) groups. None of the candidial infections were characterised as invasive. Of the potential opportunistic infections reported, there were 2 serious infections in the IFN β-1a group (strongyloidiasis and viral myocarditis) and 1 serious infection in the DAC HYP group (pulmonary tuberculosis).

Overall, this data suggests that DAC HYP might increase the risk of infection, but there is no clear signal for opportunistic infections compared to IFN β-1a.

The potential risk of progressive multifocal leukoencephalopathy (PML) due to opportunistic reactivation of the JC virus is considered separately, below.

9.6.4. Haematology

9.6.4.1. Reductions in white cells, including CD4+ and CD8+ cell counts

Phase 1 studies of DAC HYP suggested that it caused lymphopaenia, and the pivotal studies confirmed this, but changes in mean total lymphocyte counts were small. The main lymphocyte class showing a decrease in mean counts was the CD8+ subtype.

9.6.4.2. Changes in mean counts

In Study 205MS201, total white blood cell (WBC) counts were similar between the groups, but mean total lymphocyte counts were lower at Week 52 in the DAC HYP groups than the placebo group. For the combined DAC HYP groups (150 mg and 300 mg), mean WBC counts decreased 0.03% and mean lymphocyte counts decreased 5.12% from baseline to Week 52. An assessment of lymphocyte subtypes suggested an unequal effect on different cell types. Mean B-cell, CD4+ T-cell, and CD8+ T-cell counts were all lower in the DAC HYP groups, but the fall was greatest in CD8+ cells: mean B-cell counts decreased 7.70%, mean CD4+ T-cell counts decreased 7.93% and CD8+ T-cell counts decreased 47.78% from baseline to Week 52. Differences between active treatment and placebo were highly statistically significant for both CD4+ and CD9+ cells, as shown in the tables below.

In contrast to the fall in CD4+ and CD8+ cells, mean NK cell counts were higher in the DAC HYP groups than in the placebo group at Week 52, with the mean NK cell counts increasing by 49.30% from baseline to Week 52 in the combined DAC HYP groups, and no substantial difference across the two different active doses (48.14% increase in the DAC HYP 150 mg group, 50.50% in the DAC HYP 300 mg group); by contrast, there was only a 2.77% increase from baseline in the placebo group. The difference between the placebo and DAC HYP groups was highly statistically significant. The increase in NK cells with DAC HYP treatment was primarily due to an increase in the CD56bright NK subtype.

In Study 205MS301, mean WBC counts remained within normal limits and were similar between groups. There were small decreases over time in both treatment groups, with the percentage change being greater in the IFN β-1a group (< 10%) than in the DAC HYP group (≤ 5%). Mean lymphocyte counts also remained within normal limits, but showed small decreases in both treatment groups. The largest mean percentage decreases in lymphocyte count from baseline at any time point were -2.4% in the IFN β-1a group at Week 120 and -8.7% in the DAC HYP treatment group at Week 144. Changes in the mean counts of individual

lymphocyte subtypes were not clearly reported in the Summary of Clinical Safety, particularly for Study MS301. In some places, the study report mentioned CD4+ cells, and in other places it concentrated on the T-cell marker FoxP3+, which appears to identify a regulatory subset of CD4+ cells. The table excerpt below shows that FoxP3+ cells fell during treatment with DAC HYP. The patient numbers are low, indicating that lymphocyte subtypes were only assessed in a minority of subjects, as part of a pharmacodynamic assessment rather than as a major feature of safety monitoring. Summary statistics are given in Table 32, below.

Table 35. Summary statistics FoxP3+ regulatory T-cells (cells/mm^3) at Week 96, Study 205MS301

	IFN beta-1a 30 mcg	DAC HYP 150 mg
Week 96		
n	32	32
Mean	16.02	7.12
SD	11.253	6.560
Median	11.48	4.61
Min, Max	3.1, 50.7	0.6, 31.6
Percent change from baseline to Week 96		
n	28	26
Mean	0.98	-62.60
SD	80.472	21.876
Median	-14.48	-66.62
Min, Max	-83.8, 268.7	-96.0, -12.8

NOTE 1: Biomarker data collected more than 28 days after last dose date are not included.
2: PD population = subjects with at least one PD biomarker measurement.
3: Evaluable subjects = subjects with at least one measurement for the given biomarker.

In keeping with these results, the clinical study report included the following summary: *The PD response to DAC HYP seen in previous DAC HYP clinical studies was confirmed in this study and included sustained CD25 saturation, an increase in CD56bright NK cells, and a decrease in CD4+ CD127 lowFoxP3+ regulatory T-cells. CD25 saturation was apparent by Week 4, the first timepoint examined, and sustained through Week 24, the last timepoint examined. While the changes in CD4+ CD127 lowFoxP3+ regulatory T-cells plateaued by Week 12, the expansion of CD56bright NK cells appeared to plateau by Week 96. The changes in these markers were consistent with the hypothesized effect of DAC HYP on the modulation of IL-2 signalling, including decreased signalling at the high-affinity IL-2 receptor and increased signalling at the intermediate-affinity IL-2 receptor*.[5]

This summary does not mention changes to CD8+ cell counts, and a suitable summary table showing changes in CD8+ cell counts could not be located. The clinical study report mentions that CD8+ counts were similar in subjects with and without infection, but it is not clear how CD8+ counts differed across treatment groups.

The sponsor should be asked to clarify the effect of DAC HYP on lymphocyte subtypes in Study 205MS301. To some extent, the lack of clear reporting of the effect of DAC HYP on CD4+ and CD8+ cells in Study MS301 is understandable, given that the effects were already well documented in the placebo-controlled study, 205MS201. On the other hand, 205MS301 represents the only Phase 3 pivotal study in the DAC HYP study program and its duration (up to 144 weeks) potentially provided a better opportunity to study this issue than the shorter study, MS201 (52 weeks), so it would have been appropriate to provide a more detailed assessment of this issue.

9.6.4.3. Shifts in counts and values of potential clinical concern

In Study 205MS201, an analysis of shifts in counts, including values of potential clinical concern, showed a dose-dependent increase in the number of subjects with low CD4+ counts, as summarised in Table 33, below. A similar table for CD8+ cells was not provided, and the sponsor should be asked to clarify how many subjects had significant or concerning falls in CD8+ counts in Study 205MS201. This is particularly important because CD8+ cells may play a

role in preventing PML. Falls in CD8+ cells have been implicated in the mechanism by which the oral MS therapy, dimethyl fumarate, has caused PML in some subjects.

Table 36. Percentage of subjects with low post-baseline CD4 counts

Percentage of subjects with low post baseline CD4 counts

	Placebo	150 mg DAC HYP	300 mg DAC HYP	DAC Total
Number of subjects in safety population	204	208	209	417
Number of subjects in safety population with CD4 data	203 (100)	205 (100)	207 (100)	412 (100)
< 200 cells/uL	3 (1)	4 (2)	8 (4)	12 (3)
< 400 cells/uL	48 (24)	52 (25)	88 (43)	140 (34)

NOTE: Numbers in parentheses are percentages.

In Study 205MS301, potentially clinically significant changes in haematology parameters were slightly more common in the IFN β-1a group, which had a higher incidence of lymphopaenia. It was unclear how often there were significant falls in CD4+ or CD8+ cells, and this should be further clarified by the sponsor. SAEs involving low WBC cell parameters included one of agranulocytosis and one of lymphopenia, both in the DAC HYP group.

Table 37. Potentially clinically significant haematology laboratory abnormalities

Laboratory Parameters	Criterion	IFN beta-1a 30 mcg	DAC HYP 150 mg
Number of subjects in safety population		922	919
WBC Count (x10^9 cells/L)	Total n	909 (100)	913 (100)
	<3.0	53 (6)	38 (4)
	>=16	19 (2)	21 (2)
Lymphocyte (x10^9 cells/L)	Total n	909 (100)	913 (100)
	<0.8	85 (9)	70 (8)
	<0.5	18 (2)	7 (<1)
	>12	0	0
Neutrophils (x10^9 cells/L)	Total n	909 (100)	913 (100)
	<1.5	74 (8)	48 (5)
	<=1.0	10 (1)	6 (<1)
	>=12	33 (4)	32 (4)
RBC (x10^12 cells/L)	Total n	909 (100)	913 (100)
	<=3.3	9 (<1)	7 (<1)
	>=6.8	0	0
Hemoglobin (g/L)	Total n	909 (100)	913 (100)
	<=100	42 (5)	52 (6)
Platelet count (x10^9 cells/L)	Total n	909 (100)	913 (100)
	<=100	9 (<1)	10 (1)
	>=600	1 (<1)	5 (<1)

NOTE 1: Total n is the number of subjects in the safety population with at least one post-baseline value.
Numbers in parentheses are percentages.
2: Includes data up to 180 days after Last Dosing Date.

9.6.5. Progressive Multifocal Leukoencephalopathy

PML is a rare but serious inflammatory condition affecting the white matter of the central nervous system, often causing death or disability. It is caused by the JC virus, which is carried asymptomatically in a significant proportion of the population. PML is known to be a complication of JC viral activation in CNS, usually in subjects with immune compromise, such as acquired immune deficiency syndrome (AIDS) or immunosuppressive drugs.

Although the original injectable agents for MS (beta interferon and glatiramer acetate) have not been associated with an elevated risk of PML, several new MS agents do appear to cause a low

incidence of PML, particularly when associated with lymphopaenia or with reduced lymphocyte access to the CNS. In general, this risk cannot be inferred from the pivotal study data. For natalizumab, where there is a well-characterised risk of PML, the original pivotal studies did not detect the risk. Similarly, for dimethylfumarate, PML has developed in a few subjects with prolonged lymphoaenia, but this was not a feature of the original pivotal studies and only occurred during postmarketing use of the drug. PML has also been reported in subjects using fingolimod. The risk with different agents has usually been estimated at less than 1 in 1000 subjects, but varies with JC positivity and duration of exposure.

PML was not reported in subjects exposed to DAC HYP, but the studies were too brief and too small to detect a low risk of PML. In general, all agents with efficacy against MS should be considered to pose a PML risk until proven otherwise, because lymphocyte function in the CNS is important for normal immune surveillance, including protection against the JC virus, as well playing a strong role in the pathogenesis of MS. This concern is particularly relevant in view of the results of haematological monitoring, which showed that DAC HYP reduces CD4+ and CD8+ lymphocytes.

The PI should make it clear that exposure to DAC HYP has not yet been extensive enough to characterise the PML risk, and this issue should be the focus of ongoing post-marketing surveillance.

9.7. Evaluator's overall conclusions on clinical safety

Overall, the safety profile of DAC HYP has been reasonably well characterised in terms of tolerability and common side effects, but the extent to which it may cause serious idiosyncratic reactions is still unclear. Although it might be expected to pose a risk of PML, it has not yet been used in a large JC virus-positive population for long enough to characterise this risk accurately.

In terms of tolerability and common AEs, DAC HYP has an acceptable profile. In the placebo controlled study, 205MS201, the incidence of AEs was 74% in DAC HYP recipients (73% for 150 mg, 76% for 300 mg), compared to 79% in placebo recipients, as summarised in Table 35 below.

Although these percentages appear to favour DAC HYP over placebo, a direct comparison of AE incidence is unreliable because of the inclusion of MS relapses. 'MS relapse' was the most commonly reported AE, but clearly reflected efficacy rather than safety. It would have been appropriate to report total AEs excluding MS-relapse.

'Treatment-related' AEs were thought to have occurred in about 22% of subjects treated with DAC HYP. The most common AEs with an apparent causal relation to DAC HYP consisted of: injection-site pain, influenza-like illness, headache, abnormal LFTs (ALT increased, AST increased, LFT abnormal, or GGT increased), injection-site erythema or bruising, rash, eczema, nausea, lymphadenopathy, and lymphopenia. Many other common AEs seem less likely to have been causally related to treatment: nasopharyngitis, pyrexia, upper respiratory tract infection, pharyngitis, MS relapse, and fatigue. The majority of AEs were mild to moderate and responded to standard treatment or interruption or discontinuation of DAC HYP.

Skin reactions were common TEAEs and occurred in about 37% of subjects in the active-control study, 205MS301. Most of the cutaneous adverse events were mild to moderate and resolved with topical treatment or interruption of DAC HYP. About 2% of cutaneous adverse events were rated as serious. The serious cases were usually treated with systemic corticosteroids, and this should be mentioned in the PI. One severe skin hypersensitivity reaction to DAC HYP led to a patient death, albeit indirectly: the patient developed bacteraemia in the setting of an exfoliative rash, leading to the development of a psoas abscess, emboli and bowel ischaemia. It remains unclear whether this was a case of SJS.

Another death attributed to DAC HYP was a case of autoimmune hepatitis, which occurred during re-initiation of DAC HYP in a patient involved in two DAC HYP studies. This case led to more intensive monitoring in the clinical study programs. Abnormal LFTs were common in the DAC HYP studies, and were managed by interruption or discontinuation of treatment. There were no further episodes of autoimmune hepatitis, but the increased vigilance could have led to a lower incidence of severe hepatic abnormalities in this closely monitored environment than might be expected in routine clinical use. At least one DAC HYP recipient in Study MS301 satisfied Hy's Law. Other subjects had abnormal LFTs sufficient to be characterised as Hy's Law cases, but were not classified as satisfying Hy's law because alternative explanations of abnormal LFTs were considered possible. The risk of severe hepatic abnormalities has led US authorities to place a boxed warning in the US PI. It appears likely that the risk could be adequately managed in the post-marketing environment with a program of monitoring LFTs and ceasing treatment when these become sufficiently abnormal. The precise level of LFT derangement that should trigger a cessation of treatment is unclear.

Colitis was observed in some patients treated with DAC HYP. This largely resolved after DAC HYP was discontinued. The mechanism and optimal management of colitis in this setting remain unknown.

Table 38. Adverse events, DAC HYP versus placebo, Study 205MS201

	Placebo	150 mg DAC HYP	300 mg DAC HYP	DAC Total
Number of subjects in safety population	204 (100)	208 (100)	209 (100)	417 (100)
Number of subjects with an event	161 (79)	151 (73)	159 (76)	310 (74)
MULTIPLE SCLEROSIS RELAPSE	77 (38)	47 (23)	42 (20)	89 (21)
NASOPHARYNGITIS	31 (15)	30 (14)	30 (14)	60 (14)
HEADACHE	21 (10)	20 (10)	20 (10)	40 (10)
UPPER RESPIRATORY TRACT INFECTION	14 (7)	18 (9)	22 (11)	40 (10)
PHARYNGITIS	9 (4)	13 (6)	13 (6)	26 (6)
ORAL HERPES	10 (5)	10 (5)	13 (6)	23 (6)
RASH	6 (3)	12 (6)	11 (5)	23 (6)
ALANINE AMINOTRANSFERASE INCREASED	4 (2)	10 (5)	12 (6)	22 (5)
DEPRESSION	3 (1)	10 (5)	12 (6)	22 (5)
PYREXIA	2 (<1)	7 (3)	15 (7)	22 (5)
RESPIRATORY TRACT INFECTION	11 (5)	7 (3)	13 (6)	20 (5)
URINARY TRACT INFECTION	9 (4)	9 (4)	10 (5)	19 (5)
BACK PAIN	10 (5)	8 (4)	10 (5)	18 (4)
RESPIRATORY TRACT INFECTION VIRAL	5 (2)	8 (4)	10 (5)	18 (4)
INFLUENZA	11 (5)	5 (2)	12 (6)	17 (4)
DIARRHOEA	4 (2)	7 (3)	8 (4)	15 (4)
RHINITIS	3 (1)	9 (4)	6 (3)	15 (4)
FATIGUE	10 (5)	6 (3)	8 (4)	14 (3)
ASPARTATE AMINOTRANSFERASE INCREASED	2 (<1)	7 (3)	6 (3)	13 (3)
VIRAL INFECTION	4 (2)	4 (2)	9 (4)	13 (3)
BRONCHITIS	5 (2)	4 (2)	8 (4)	12 (3)
NAUSEA	2 (<1)	4 (2)	8 (4)	12 (3)
INFLUENZA LIKE ILLNESS	6 (3)	6 (3)	5 (2)	11 (3)
OROPHARYNGEAL PAIN	4 (2)	6 (3)	5 (2)	11 (3)
ANAEMIA	1 (<1)	6 (3)	4 (2)	10 (2)
PARAESTHESIA	10 (5)	3 (1)	6 (3)	9 (2)
DIZZINESS	5 (2)	2 (<1)	6 (3)	8 (2)
ACNE	1 (<1)	1 (<1)	6 (3)	7 (2)

NOTE 1: Numbers in parentheses are percentages.
2: A subject was counted only once within each preferred term.
3: Preferred terms are presented by decreasing incidence in the total column.

An excess of mild to moderate depression was observed in subjects treated with DAC HYP, compared to placebo, in Study MS201. The incidence of depression appeared to be similar to that observed with IFN β-1a in Study MS301. DAC HYP should be contraindicated in patients with a recent history of severe depression.

The use of DAC HYP was associated with a significant reduction in CD4+ and CD8+ T-cells, which might be expected to increase the risk of PML. Experience with other disease-modifying

agents in MS suggests that this risk will not be fully characterised until the drug has been used in a large population of at-risk, JC-positive subjects.

The pivotal studies showed a mild excess of infections in DAC HYP recipients, but a substantially increased risk of opportunistic infections, was not observed. The risk of infections should remain a focus of post marketing surveillance.

In conclusion, the tolerability of DAC HYP is broadly acceptable. In terms of serious but rare safety issues, DAC HYP appears to be associated with a risk of severe reactions in a small proportion of subjects. These include:

- hepatic reactions, including autoimmune hepatitis
- hypersensitivity reactions, incuding skin reactions and anaphylaxis
- lymphopaenia, especially affecting CD4+ and CD8+ lymphocytes
- a theoretical risk of progressive multifocal leukoencephalopathy.

10. First round benefit-risk assessment

10.1. First round assessment of benefits

The benefits of DAC HYP in the proposed usage are:

- DAC HYP appears to reduce annualized relapse rate by about 50 to 54%, relative to placebo ($p \leq 0.0002$)
- DAC HYP reduces relapse rate by about 45%, relative to once-weekly IFN β-1a ($p < 0.0001$)
- DAC HYP reduces the proportion of subjects relapsing by 44 to 47%, relative to placebo, and by 34 to 35%, relative to IFN β-1a, depending on the duration of follow-up. (Note that this is less benefit than claimed by the sponsor)
- DAC HYP reduces radiological evidence of disease activity, including the reduction of new Gd-enhancing lesions by 69 to 78%, and new or newly enlarging T2 lesions by 70 to 79%, relative to placebo ($p < 0.0001$)
- DAC HYP is associated with a strong trend to reduced disability progression
- DAC HYP produced a broadly similar benefit across all major subgroups in the study population.
- DAC HYP has not been studied in subjects with overt SPMS, and its efficacy in this population is unknown.

10.2. First round assessment of risks

The risks of DAC HYP in the proposed usage are:

- a high incidence of skin reactions (about 37%, with about 2% rated as serious)
- hepatic reactions, including potentially severe or fatal autoimmune hepatitis
- hypersensitivity reactions, including anaphylaxis
- lymphopaenia, especially affecting CD4+ and CD8+ lymphocytes
- a theoretical risk of progressive multifocal leukoencephalopathy.

10.3. First round assessment of benefit-risk balance

DAC HYP reduces relapse rate in subjects with RRMS, but its use is associated with significant safety concerns. The efficacy of DAC HYP appears to be broadly comparable to other new disease-modifying agents, in terms of reducing relapse rate in subjects with RRMS, so it needs to be considered alongside those other agents. Like most other disease-modifying agents at the time of their registration, DAC HYP has not produced clear benefits in terms of reducing disease progression, but it is expected to reduce the accumulation of disability by preventing overt clinical relapses as well as new plaques evident on MRI. The submitted evidence suggests that a benefit on progression is very likely, but robust statistical proof is still lacking. Despite this, a benefit in terms of reducing relapse rate is a worthwhile clinical achievement in its own right, even without a proven benefit on progression, and one that would be attractive to patients and clinicians, if that reduction in relapse rate could be delivered with acceptable risk, relative to other available agents. Whether the observed reduction in relapse rate outweighs the safety concerns will depend on the extent to which the individual patient considering treatment is at risk of further relapses (and at risk of disability related to those relapses).

Compared to the first generation disease-modifying agents, such as beta interferon and glatiramer acetate, DAC HYP does not offer the same relatively benign safety profile. Although beta interferons have been associated with a number of tolerability concerns and can cause abnormalities of liver function tests, the risk of severe reactions (including severe derangements of liver function) appears higher with DAC HYP. The risk of skin reactions also appears high, with 2% of subjects experiencing serious skin reactions that led to use of systemic steroids. Like other monoclonal antibody preparations, DAC HYP may also cause acute hypersensitivity reactions and poses a risk of anaphylaxis. It is likely to increase the risk of PML, but this remains unclear.

The efficacy of DAC HYP is clearly superior to once-weekly IFN β-1a, which might justify increased safety risks, but DAC HYP has not been directly compared to more frequently administered beta interferon which is widely believed to be more effective than once-weekly IFN β-1a and has proven to be superior to IFN β-1a in head-to-head comparisons. The benefit of DAC HYP against more effective interferon regimens is likely to be minor, meaning that a substantial safety risk may not be justified.

For subjects with highly active disease, and particularly for subjects with a proven failure of beta interferon therapy, a low risk of serious complications is likely to be considered acceptable when choosing a new disease-modifying agent. Balanced against the high likelihood of frequent relapses, progressive motor disability, sensory disturbances and cognitive decline in the absence of an effective MS treatment, the rare occurrence of hepatic reactions and other serious complications carries less weight. If DAC HYP were known to reduce disability progression, patients would be expected to accept significant safety risks, but unfortunately there is no robust confirmation of this at present. The fact that DAC HYP reduces relapse rate by at least 50%, coupled with the fact that a large proportion of disability progression is known to come from incomplete recovery from relapses, suggests that DAC HYP could have a useful role in subjects with a high risk of relapses. Provided that the risks and benefits are made clear to patients and clinicians, they are in the best position to decide what risk they are prepared to accept to achieve a 50% reduction in relapse rate, and whether DAC HYP is an appropriate choice compared to other available agents. None of the new agents is without some significant safety concerns and some patients show poor tolerability of other new agents, such as dimethylfumarate, so it is expected that DAC HYP will find a use in some patients.

Like most other disease-modifying agents, DAC HYP has not been tested in subjects with overt SPMS, so the benefit-risk profile in this group is unknown. Immune-modifying agents have generally been less effective in subjects with progressive disease, and the same is expected to be true of DAC HYP. The supplementary evaluator was not convinced by the sponsor's argument that, despite clear entry criteria that excluded SPMS, the pivotal studies inadvertently included

some SPMS subjects, and this inadvertent inclusion therefore justifies use of DAC HYP in the broader population of subjects with SPMS. Only a study that explicitly focussed on SPMS subjects could demonstrate efficacy in this group with sufficient clarity that a rational decision could be made about the benefit-risk profile in SPMS.

In conclusion, the benefit-risk balance of DAC HYP for the proposed indication, which includes all forms of relapsing MS, is not known to be favourable. There is not sufficient evidence to recommend the use of DAC HYP in subjects with SPMS and the proposed indication does not match the entry criteria of the pivotal studies.

The benefit-risk balance for DAC HYP for a modified indication is expected to be favourable, if DAC HYP is used exclusively in subjects with RRMS, who are still experiencing relapses (or who are avoiding relapses by use of an alternative disease-modifying agent), who accept the risks, and who can receive DAC HYP in a closely monitored prescribing environment.

11. First round recommendation regarding authorisation

The sponsor's application to register DAC HYP for all subjects with relapsing forms of MS should be rejected.

Authorisation should be reconsidered after the sponsor has:

- provided adequate answers to the clinical questions raised by the evaluator
- addressed concerns raised about the proposed PI
- provided a satisfactory mechanism to ensure DAC HYP is only prescribed by clinicians aware of its safety issues, with appropriate monitoring of LFTs
- modified the wording of the indication so that it matches the study population in the two pivotal studies.

11.1. Issues raised in the first round clinical evaluation report

The First Round clinical evaluator expressed many concerns about the proposed PI that are shared by the supplementary evaluator. These are addressed below.

11.1.1. Efficacy

11.1.1.1. First round concerns

'In the narrative under Table 2 reference is made to the MSIS-29 endpoint for Study 205MS201 under this heading and efficacy results and a p-value are given. However the difference for DAC HYP 150 mg versus placebo was not considered statistically significant per the sequential closed testing procedure because the procedure required that the 300 mg dose group be tested first and achieve statistical significance before the 150 mg dose group could be tested. The result should be amended to reflect this or removed.'

Supplementary evaluator's response

The supplementary evaluator fully agreed with this criticism.

11.1.2. Indications

11.1.2.1. First round concerns

'The indication, treatment of relapsing forms of MS, is considered too broad and should be amended to reflect the target patient population and the primary endpoint investigated in the clinical studies. For example:

- *'DAC HYP is indicated in patients aged 18-years or over who have RRMS who have had two or more clinical relapses within the previous 3 years with at least 1 clinical relapse in the 12 months prior to treatment';* or
- *'One or more clinical relapses and 1 or more new MRI lesions (Gd-enhancing and/or T2 hyper-intense lesion) within the previous 2 years, with at least one of these events in the 12 months prior to treatment.'*

Supplementary evaluator's response

The supplementary evaluator agrees with the statement that the proposed indication is too broad. A strict interpretation of the pivotal studies suggests that efficacy for DAC HYP has only been demonstrated in subjects with RRMS and evidence of recent disease activity, as evidenced by at least 1 relapse in the previous year or 2 relapses in the previous 3 years (or a radiological substitute for clinical relapses). The supplementary evaluator does not fully agree that this highly restricted definition of the target population should be used in the PI.

Firstly, there is a spectrum of disease patterns in MS ranging from those with mild disease to those with more severe disease, and another spectrum from those primarily affected by relapses to those primarily affected by progression. Efficacy is likely to extend along both spectra to provide some benefit in subjects who are slightly outside the strict entry criteria applied to the pivotal studies.

Secondly, and more importantly, there is a large section of the MS population who have already been on disease-modifying treatment, prior to considering a switch to an agent like DAC HYP, and who are likely to have experienced a reduction in relapse rate because of this treatment, relative to what they would have experienced if they had not been treated. This reduction in relapse rate could mean that they have not experienced two relapses in the previous 3 years solely because they respond well to immunomodulatory treatment. If such subjects are switching disease-modifying agents because of poor tolerability with an existing agent, it would be unfair to exclude them on the basis of their good response to the previous agent – subjects showing a good response may be the most appropriate subjects to continue a disease-modifying agent. Conversely, if they are switching because of poor efficacy of the previous agent, as evidenced by clinical and radiological evidence of ongoing relapses, then they are likely to satisfy the entry criteria for the pivotal studies anyway, or at least be broadly similar to the study population.

Treatment-naïve RRMS subjects without recent relapses were not eligible the pivotal studies, and cannot be considered to have a low relapse rate that is attributable to treatment, so it would be reasonable to exclude this category from the proposed indication. One problem is that the indication could become unwieldy if it attempted to cover every clinical possibility. For instance, the indication could be expressed as follows:

'DAC HYP is indicated in patients with Relapsing and Remitting MS (RRMS) aged ≥ 18 years who have had ≥ 2 clinical relapses within the previous 3 years OR ≥ 1 clinical relapse and ≥ 1 new MRI lesion within the previous 2 years; or are switching from a different disease-modifying treatment. DAC HYP is not indicated in subjects with Secondary Progressive MS or Primary Progressive MS'

This wording is somewhat simplified compared to the entry criteria for the pivotal studies (it does not require that subjects have had a relapse in the previous 12 months, for instance), but it allows for the possibility of subjects switching treatments.

Given the safety profile of DAC HYP, it would also be reasonable to restrict treatment to subjects who have failed on other disease-modifying agents. The US PI recommends this, without formally insisting on it:

'Zinbryta is an interleukin-2 receptor blocking antibody indicated for the treatment of adult patients with relapsing forms of multiple sclerosis (MS). Because of its safety profile, the use of

Zinbryta should generally be reserved for patients who have had an inadequate response to two or more drugs indicated for the treatment of MS.'

To some extent, clinicians may be best placed to balance the risk and benefits of switching agents, and it is probably sufficient to restrict the indication to subjects with RRMS, while warning clinicians not to prescribe DAC HYP unless they feel the benefits outweigh the risks. This would lead to a simpler indication, as follows:

'*DAC HYP is indicated in patients with Relapsing and Remitting MS (RRMS) aged ≥ 18-years who have ongoing relapses; or are switching from a different disease-modifying treatment. DAC HYP is not indicated in subjects with Secondary Progressive MS or Primary Progressive MS.'*

On balance, this is the preferred wording for the supplementary Evaluator, but the earlier wording requiring evidence of a sufficient number of relapses would also be acceptable.

These comments in favour of simplification should not be taken as supportive of a further relaxation of the definition of the target population, to include subjects who do not even have RRMS, because these subjects were not the focus of the pivotal studies and there has been ample evidence over decades that efficacy in RRMS does not translate well to subjects with SPMS.

11.1.3. Contraindications

11.1.3.1. First round concerns on hepatic impairment

'Consideration should be given to adding a contraindication for patients with hepatic impairment as these patients were excluded from the clinical studies. Since the safety profile for DAC HYP includes elevations in LFTs and it is not known how patients with pre-existing hepatic impairment will be affected.'

Supplementary evaluator's response

The supplementary evaluator supports listing hepatic impairment as a contraindication to use of DAC HYP. Currently, hepatic impairment is only listed as a precaution, and only if the pre-existing impairment is severe:

'*Zinbryta is not recommended for use in patients with pre-existing severe hepatic impairment.'*

Hepatic abnormalities (and one liver-related fatality) were observed in the pivotal studies despite the exclusion of subjects with pre-existing hepatic impairment. As noted by the First Round clinical evaluator, hepatic complications might be more frequent if the drug was used in such subjects.

11.1.3.2. First round concerns on concomitant use of immunosuppressant therapy

'Since this product is an immunomodulator, a contraindication for use in patients with increased risk for opportunistic infections [should be added to the PI], including those immunocompromised due to current or recent immunosuppressive therapies or systemic medical conditions resulting in significantly compromised immune system function (for example human immunodeficiency virus, organ transplant, active malignancy); subjects such as these were also excluded from the clinical studies.'

Supplementary evaluator's response

The supplementary evaluator agrees that clinicians should avoid combining DAC HYP with other immunosuppressive treatments or conditions.

11.1.3.3. First round concerns on use in patients with a history of depression

'A contraindication for patients with a recent history of severe depression should be added.'

Supplementary evaluator's comment

This seems appropriate.

11.1.4. Precautions

11.1.4.1. First round concerns restricting prescribing to neurologists only

'Although it is unlikely that any physician other than a neurologist will start treatment with DAC HYP, wording should be added to state that: DAC HYP should only be initiated by neurologists experienced at treating and monitoring patients who have MS.'

Supplementary evaluator's response

Arguably, this statement should be even more strongly worded. The US PI restricts use of DAC HYP to treatment within the context of a strict 'Risk Evaluation and Mitigation Strategy (REMS) called the Zinbryta REMS Program'. In Australia, a similar program should be considered. There is a clear precedent. When the monoclonal antibody natalizumab was registered for treatment of MS in Australia, it was initially restricted to neurologists who had undergone training about the particular risks associated with natalizumab, including PML. Balancing the benefits and risks of DAC HYP will not be straightforward, given that other agents offer similar efficacy, and it would be reasonable to restrict DAC HYP to neurologists with experience in MS who have undergone specific training about the risks of DAC HYP.

11.1.4.2. First round concerns regarding PML risk

'The patient population in the clinical studies was not large enough to detect cases of progressive multifocal leukoencephalopathy. The possibility of development [of PML] cannot be excluded and this should be reflected in this section. For example: *'Although not seen in the clinical development program for DAC HYP, due to the immunomodulatory action, the development of PML remains a rare possibility. Patients should be screened for JC virus prior to initiation of treatment to minimise the risk.'*

Supplementary evaluator's response

The proposed addition to the PI is appropriate, and the supplementary evaluator agrees that JC serological status should be monitored in subjects on DAC HYP. There is currently insufficient evidence to justify a policy of restricting DAC HYP to JC-negative subjects, but knowledge of JC status could allow a heightened index of suspicion for PML, more frequent MRI monitoring, and more appropriate action if PML cases are reported internationally. The following wording is proposed:

'Although progressive multifocal leukoencephalopathy (PML) was not seen in the clinical development program for DAC HYP, the development of PML remains a possibility, because of the immunomodulatory action of DAC HYP. Patients should be screened for JC virus prior to initiation of treatment, and JC-positive subjects should be monitored closely, to minimise the risk.'

11.1.4.3. First round concerns regarding depression risk

'Since an excess of depression was seen in subjects who received DAC HYP in the placebo controlled study appropriate warnings should be placed in this section. For example: *In a placebo controlled study depression was seen more commonly in subjects who received daclizumab. Patients should be warned of the possibility of experiencing depression and should seek medical advice if they have alterations in mood.'*

Supplementary evaluator's response

The proposed addition to the PI is broadly appropriate. For consistency with the rest of the PI, the placebo-controlled study should be named.

12. Clinical questions

12.1. Additional expert input

The risks of serious hepatic injury with DAC HYP may be worse in subjects with pre-existing hepatic disorders. The opinion of a hepatologist should be sought about the degree to which pre-existing abnormalities of liver function should be considered a contraindication to DAC HYP.

Consideration should be given to obtaining expert advice about the PML risks likely to be associated with DAC HYP. Alternatively, it could be assumed that the risks are significant, on the basis of the observed effects of DAC HYP on lymphocytes and the experience with other disease-modifying agents in MS.

12.2. Efficacy

Q1. If the question had not already been raised by the EMA, it would be important to ask the sponsor to justify the proposed indication for 'relapsing forms of MS' given that the pivotal studies of DAC HYP were specifically conducted in subjects with RRMS. The sponsor's response to this issue has already been discussed in detail in Section: Supplementary material submitted to the EMA in relation to pivotal efficacy data and largely rejected. The sponsor is invited to read this section and suggest a more appropriate indication that matches the entry criteria of the pivotal studies.

Q2. What evidence is available to establish the minimum effective dose of DAC HYP? Given that Study 205MS201 showed no substantial difference between 150 mg and 300 mg, it appears possible that lower doses, such as 100 mg SC 4-weekly, would also have acceptable efficacy, and these lower doses might have improved safety. Please comment on the proposed dose in light of these observations.

Q3. In the efficacy results for Study 205MS201, the precise meaning of the p-values cited next to footnote '(b)' in the sponsor's table was not clear (please see Table 6 of this document). The sponsor should be asked to clarify what test has been used and what the p-values indicate.

Q4. Please provide clear summary tables of the number and nature of protocol deviations in the two pivotal studies. In Study 205MS201, a clear tale grouping protocol deviations by type and severity was not provided. In Study 205MS301, major protocol deviations were summarised by the sponsor as follows:

> *'Overall, the incidence and category of major protocol deviations were similar between the two treatment groups. The most common major deviations (≥ 20%) were 'Informed Consent' (32% IFN β-1a versus 33% DAC HYP), 'Key Study Procedures' (28% IFN β-1a versus 27% DAC HYP), and 'Other' (27% each group).'*

The sponsor's text provided a link to a table of protocol deviations, reproduced below, but this table lacked detail and it is not possible to determine whether the deviations related to 'Key Study Procedures' or 'Other' substantially compromised the study. The sponsor should clarify the nature of the deviations and consider the extent to which these deviations could have compromised the study.

Table 39. Summary of Major Protocol Deviations, Study 205MS301

Summary of major protocol deviations

	IFN beta-1a 30 mcg	DAC HYP 150 mg
Number of subjects in the ITT population	922 (100)	919 (100)
Informed Consent	294 (32)	299 (33)
Eligibility	25 (3)	29 (3)
Study Tx Administration	144 (16)	148 (16)
Prohibited Con Med	44 (5)	52 (6)
Key Study Procedure	262 (28)	244 (27)
Other	253 (27)	244 (27)

NOTE: Numbers in parentheses are percentages.

Q5. Did the sponsor attempt to identify the extent of inadvertent unblinding in the pivotal studies, because of tell-tale side effects? If not, why not? Please estimate the extent of unblinding by considering the incidence of side effects in the pivotal studies.

12.3. Safety

Q6. (A) For lymphocyte subtypes, such as CD4+ and CD8+ T-cells, changes in mean counts and the incidence of clinically significant shifts were not always clearly reported, particularly for the active-controlled study, Study 205MS301. The sponsor should be asked to clarify the effect of DAC HYP on lymphocyte subtypes in Study 205MS301, providing brief summary tables for changes in mean counts and the incidence of potentially clinically significant shifts in counts.

Q6. (B) In Study 205MS201, an analysis of shifts in counts, including values of potential clinical concern, showed a dose-dependent increase in the number of subjects with low CD4+ counts, but a similar table for CD8+ cells was not provided. The sponsor should clarify how many subjects had significant or concerning falls in CD8+ counts in Study 205MS201.

Q7. The sponsor should provide estimates of the risk of DAC HYP causing PML in JC positive subjects, and comment on the extent to which baseline lymphopaenia could modify this risk.

13. Second round evaluation of clinical data submitted in response to questions

The secondary clinical evaluation and the questions posed in Section 12 above were taken into consideration by the Delegate when writing the Delegate's Overview (see Overall conclusion and risk/benefit assessment in AusPAR).

14. References

Durelli L et al; Independent Comparison of Interferon (INCOMIN) Trial Study Group. Every-other-day interferon beta-1b versus once-weekly interferon beta-1a for multiple sclerosis: results of a 2-year prospective randomised multicentre study (INCOMIN). Lancet. 2002 Apr 27;359(9316):1453-60.

Gold R et al; DEFINE Study Investigators. Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis. N Engl J Med. 2012 Sep 20;367(12):1098-107.

Schwid S et al; Full results of the Evidence of Interferon Dose-Response-European North American Comparative Efficacy (EVIDENCE) study: a multicenter, randomized, assessor-

blinded comparison of low-dose weekly versus high-dose, high-frequency interferon beta-1a for relapsing multiple sclerosis. Clin Ther. 2007;29(9):2031-2048.

Davies H et al; When can odds ratios mislead? BMJ. 1998;316:989–991.

Deeks, J; When can odds ratios mislead? Odds ratios should be used only in case-control studies and logistic regression analyses BMJ. 1998 Oct 24; 317(7166): 1155. PMCID: PMC1114127

Therapeutic Goods Administration

PO Box 100 Woden ACT 2606 Australia
Email: info@tga.gov.au Phone: 1800 020 653 Fax: 02 6232 8605
https://www.tga.gov.au