Author’s response to reviews

Title: Evaluation of neurological changes in secondary progressive multiple sclerosis patients treated with immune modulator MIS416: Results from a Feasibility Study

Authors:

Gill Webster (gill@innateimmuno.com)
Dalice Sim (dalice.sim@otago.ac.nz)
Anne LaFlamme (Anne.LaFlamme@vuw.ac.nz)
Nancy Mayo (nancy.mayo@mcgill.ca)

Version: 1 Date: 01 Sep 2017

Author’s response to reviews:

Thank you for the thoughtful comments on our study. We address each of the reviewers’ comments below. Substantial changes made to the text are highlighted in yellow in the main text.

Reviewer #1

Q: The background barely touches on the disease itself but focuses more on therapies and outcomes. It is customary to let the reader know why it is important to study this topic. It would have been easy to explain a small sample size if the background indicated how rare the condition is.

R: The background section has been re-written to inform the reader on the distinct and challenging aspects of treating and monitoring secondary progressive MS, highlighting the requirement for additional instruments that measure disease activity. The remainder of the abstract has been edited to keep the word count within the limit.

Q: I would suggest revising the last sentence on the background. Replace "support" with "investigate".

A: As suggested, this has been changed.
Q: It is stated that the study was conducted in New Zealand, but the context and setting of the trial are not reported in sufficient detail. Where were the patients recruited from? What were the exclusion/inclusion criteria? Was there a target sample size? What was the sampling strategy?

A: Further details describing the purpose and design of the clinical trial as well as the patients have been added to the methods section.

Q: Please provide more details on how the continuous data from multiple measures was collapsed into a binary variable. Did a patient have to be a responder on all measures to be classified as responder overall?

A: The text stipulates that minimal important change (MIC) was used to define a positive patient responder status as estimated from published studies.

For EDSS, MIC was 0.5 [27].

From the MFSC [24], MIC values were determined for the following performance rated outcomes (PerfOs): Gait Speed (GS) 0.10m/sec [28];

Paced Auditory Serial Addition Test (PASAT) 9 (based on changes greater than estimates of the practice effect) [29, 30];

Nine Hole Peg Test (NHPT) 20% [29].

For the SF-36 subscales, a change in 10 points on a 0-100 scale for each score was considered a MIC [32], (equivalent to 1/2 standard deviation (SD) as SD is approximately 20 in adults >55 years from a large nominative sample) [30].

The text also stipulates that each patient was classified as a responder or non-responder on the EDSS and the 3 PerfOs and 4 PROs from the SF-36. Response across measures was summarized using the total number of measures with an observed response. Participants were ranked in order of total responses across PerfOs and PROs.
As this was already stipulated in the methods section, no additional text was added.

Q: Under analysis of immunological parameters the maximum recorder (recorded) response across all time points were analysed. What are these timepoints?

A: The time points are stipulated in the preceding methods section “Quantification of plasma MIS416 immune biomarkers analysis”. They are also stated in the abstract, and have been added into the introduction.

Q: Under the analysis of immune biomarkers it is stated that patients were grouped into high, medium or low responder groups. How were these categories determined? This appears in the results, but the approach used should be described in the methods section.

A: The number of outcomes with positive response were summed across the PerfOs and PROs and the subjects were ranked by the total. Ties were broken by the number of PerfOs of response, and if that was 0, by the number of PROs of response. This information was included already in the methods section.

Q: Were these rankings planned a-priori or determined post-hoc?

A: The categories were determined post hoc as there were quite clear groupings based on the total number of responsive outcomes: >3; 2, 0 or 1. As this was a feasibility study, in particular, feasibility of using and alternate approach the analysis of multiple outcomes, we could not do too much a priori planning.

Q: Please report the genders of the participants.

A: The gender of the trial participants has been added to the methods section “Patients”.

Q: In this section "Concordance of responder ranking with MIS416 pharmacodynamic immune response" please note that it is not usual to have background/discussion material in the results section. This section is reserved for the findings from this particular study and not others. You can use the methods section to provide context for your results.

A: This section has been re-written to address this point.
Reviewer #2:

Q: My initial understanding of the problem outlined in this paper was that the Expanded Disability Status Scale (EDSS) was not very sensitive to change for people with severe secondary progressive multiple sclerosis and so the authors were investigating using a new measure that they hope to be introduced in the literature as an alternative. They take pre post patient reported outcome scores and performance rated outcomes scores as well as a number of biomarkers.

A: In fact, we were not investigating a new measure, we were testing the feasibility of using different PerfOs and PROs as endpoints recognizing that no one outcome is likely suitable to reflect the impact of a drug targeting the wide variety of disabilities experienced by people with secondary progressive MS. We added information on this method in the paper.

Q: The authors state that they have used the CONSORT to report findings but I found myself having to go backwards and forwards between the methods, results and the tables to work out things like which exact measurements were being used, what they meant by clinical responder status and immunological response. It would help the reader to have all the outcomes and definitions of responder described in the measurements section.

A: The updated methods section has hopefully clarified the measurements and how they were applied.

Q: The Fatigue Severity Scale is mentioned early on but then is not mentioned again.

A: The FSS was originally included in the test battery but we did not want to include two PROs both measuring fatigue and as the other PROs came from the SF-36, to add homogeneity, we included only SF-36 PROs. Of the 11 persons, only 2 had discordance responses on FSS and MHI: both showed responses on MHI but one of these showed no change on the FSS and one showed an increase in fatigue on the FSS after but this person showed an abnormally low level of fatigue at baseline. In sum, 9 of the 11 had concordant responses indicating the wisdom of choosing only one. In addition, the items of the FSS (n=9) do not measure severity of fatigue but rather causes, consequences and impact on daily life. We clarified this in the text.

Q: Then there are different numbers of subgroups from the SF 36 quoted as being of interested: 4 early on but all 8 in the results table.

A: The SF-36 has 8 subscales which we reported on in Table 1 to describe the sample. Of these, we chose the 4 most closely related to the biological action on the drug under study: physical...
function to reflect the everyday impact of improved gait speed or walking; mental health to reflect the impact on depression either primary effect or secondary from walking better; vitality for fatigue as it is the most distressing symptom of people with MS and not at all captured by the EDSS; and final general health as improved walking, mood, and fatigue will impact this outcome and it is an important patient centered outcomes. The other 4 subscales, pain and the 3 role subscales were not included as pain is often either neuropathic in origin or musculo-skeletal secondary to abnormal walking pattern and the role variables are immediate downstream outcomes from improved mobility and mental health. We clarified this in the text.

Q: It was good to see MICs reported and that section was clear. This section could also have stated what the 4 (or 8) PROs were too.

A: Yes, addressed above and added to the text.

Q: With regard to the patient's immunological response profile to MIS416 therapy, they describe that for the analysis they used the maximum-recorded response across all time points but later on in the results it looks like they used the maximum post scores from time point 1-4 which would make more sense.

A: This has been addressed above (reviewer #1) and the text in methods “Quantification of plasma MIS416 immune biomarkers analysis” and “Statistical analysis of immunological parameters” has been clarified.

Q: They talk about normalizing the data but then don't explain how they normalised it.

A: Text explaining normalization of immune factor levels has been added.

Q: They describe immune factors/parameters and resemblances between individuals to identify clusters of patients. MDS needs writing in full.

A: The text has been adjusted.

Q: I am not clear how the authors define ‘immunological response’. For the PROs and the PerFOs it was a 0/1 response but I am not sure if it's a 0/1 for immunological Response
A: The definition of an immunological response is based on an increase in plasma levels of one or more immune factors from baseline, pre-treatment levels. This definition has been added to the text.

Q: Then are they comparing these for the 4 PROs only? A cluster of size n=2 is also very small to draw any conclusions from

A: No conclusions were drawn from these clusters, they are described only.

Q: Biomarker levels for patients were then compared across groups according to their clinical responder status where n=5 high, n=3 medium and n=3 low and then an ANOVA was applied. It is not clear where the cut offs for high, medium and low have come from. However, I feel that the group sizes are too small for these analyses

Upon revision, the responder status has been adjusted in table 4, where n=4 high, n=4 medium and n=3 low. The basis for grouping into high, medium and low clinical responders has been explained above, in response to reviewer #1. The patient numbers in each group are small, however it is acceptable to use their immune response data for an exploratory post-hoc analysis such as this, which aims to find patterns between the subgroups which can then be used to formulate a-priori hypothesis testing studies.

Q: Some of the detail from line 250 I think describe methodology rather than results

A: This section has been re-written.

Q: I find the discussion section unclear also, what are they aiming for?

A: Discussion was redrafted

Q: How are they planning to determine what a good response is and what is better than the EDSS score that they started out. Are they claiming that 5 people responding to change on one outcome (EDSS) is worse than 7 people responding to change across 4 Perfos?

A: The MS world is moving away from the EDSS as a primary outcome (LaRocca NG, Hudson LD, Rudick R, Amtmann D, Balcer L, Benedict R, Bermel R, Chang I, Chiaravalloti ND, Chin P, Cohen JA, Cutter GR, Davis MD, DeLuca J, Feys P, Francis G, Goldman MD, Hartley E,

The outcomes for this feasibility study reflect this movement away, but with multiple outcomes, there is a statistical challenge in providing a valid estimate. This is even more challenging in small sample sizes, and the methods shown in this study present one solution to what will be a frequent problem, especially in pilot studies. A discussion about future outcomes for MS trials is beyond the scope of this study.

Q: Have ceiling effects been thought about?

A: The people in this study have major disability and have values on all outcomes well below norms. For example, the norms for the SF-36 have been provided to illustrate the extent of the disability. Hence, ceiling effects are not an issue here.

Q: How are they deciding that they can proceed to a larger study?

A: A clinical response on 4 or more of the 11 participants would be unlikely under the null hypothesis of 1 responder by chance. Here 7 responded on at least 1 PerfO and 8 responded on at least 1 PRO. We therefore believe a larger study is warranted, to formally test these patient centered outcomes, over a longer period of treatment.

Q: They describe including more PROs in the next study, but then how would they choose between them?

A: For regulatory purposes, a standard outcome needs to be used – this will be the MS Functional Composite. Use of this measure requires large sample sizes and comprises some of the measures used here (gait speed and manual dexterity). The other outcomes will be used for secondary analyses for specific motor outcomes in which case it is best to have more than one outcome to reflect the complex impact of MS on motor impairment and subsequently on activity and participation. A discussion of this is beyond the scope of this article.

Q: It would be helpful to set up criteria of what a new and better outcome should be - they mention sensitivity to change and MIC on some outcomes, but what about on the biomarkers? I
think it would be helpful to concentrate on looking at the potential new PROs and describing the level of change and sensitivity in relation to the bio markers or perfos that are deemed most important and then have clear criteria relating to the process of carrying out this piece of research as well as determining if what they have found means the feasibility is a success or not.

A:Biomarkers that have been evaluated in this study relate to MIS416 activity, as opposed to disease activity. In the context of this feasibility study, they have been examined to explore if there is any relationship between the extent of the patient immune response to the treatment the overall clinical change, rather than any single outcome measure. There is no single biomarker that has been identified to date, that is a biomarker for MS disease activity overall or any of the disability measures. We currently speculate that the immune biomarkers could be useful to establish a reference range for certain induced immune factors. This could then be used to determine whether the patient is hyper-responsive to the treatment or not. This could translate into a high or low therapeutic dose being recommended for any individual patient.

Reviewer #3:

Q:The limitations of the study (mainly the lack of a control group and small sample size) are well acknowledged. However, I did find it a bit difficult to read. Where possible, I recommend the phrases to be shorter (i.e. Lines 88-92, 109-112, 112-117). I recommend that it should be stated right from the beginning, in a clearer manner, which are the new outcomes proposed and which are the "classic" ones

Sentences have been grammatically re-phrased/re-written where possible.

A:There were no new outcomes proposed in this study, rather this was an exploratory study geared around evaluating outcomes from commonly used, neurological assessment tools in addition to the classic EDSS. The outcomes evaluated are detailed clearly in the method section.

Q:In the methods section I recommend to clearly state inclusion/exclusion criteria for the pilot study (not just regarding EDSS but all the criteria).

A:This information has been added
Q: In the abstract, I recommend to state the clear endpoints for the new study, both the immunological ones and the clinical ones).

A: Since this was a feasibility study, the objective was to conduct a meaningful post-hoc evaluation of a range of patient clinical measures, to determine whether they had potential usefulness or not, for evaluating clinical status in a future study. For the immunological response patterns, the objective was to use them to measure the patient response to treatment. We think these points in the abstract are clear.

Note that the abstract has been re-drafted.

Q: I recommend starting the discussion with three sub-titles: Immune response, Clinical - patient reported outcomes, and the third would be the relationship between the two

A: The primary objective of this feasibility study was to evaluate patient centered outcomes as additional assessment tools alongside EDSS for determining the clinical status of SPMS patients, based on post-hoc analysis of a safety and tolerability study- accordingly we think these data should be discussed first. The immune response was explored as a secondary objective.

The remainder of the discussion has been re-ordered according to the recommended sub titles.