21 February 2013
EMA/608526/2012
Committee for Medicinal Products for Human Use (CHMP)

Overview of comments received on 'Guideline on similar biological medicinal products containing interferon beta' (EMA/CHMP/BMWP/652000/2010)

Interested parties (organisations or individuals) that commented on the draft document as released for consultation.

Stakeholder no.	Name of organisation or individual
1	Synthon BV
2	EBE (European Biopharmaceutical Enterprises)
3	Bayer Pharma AG
4	Medicines Evaluation Board (The Netherlands)

7 Westferry Circus ● Canary Wharf ● London E14 4HB ● United Kingdom
Telephone +44 (0)20 7418 8400 **Facsimile** +44 (0)20 7418 8416
E-mail info@ema.europa.eu **Website** www.ema.europa.eu

1. General comments – overview

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
1	**Which data sets are preferred for which analysis?** The guideline recommends 2 analyses; one for assay sensitivity and one for equivalence. Could the CHMP give guidance on the preferred data sets for analyses. I.e. should the data set obtained Per Protocol (PP) be used for testing assay sensitivity, whereas the data obtained by Intention To Treat (ITT) can best be used to test equivalence?	The principles for statistical analysis of superiority and equivalence trials should be followed. For assay sensitivity, testing for superiority should be performed in the ITT data set. Testing for equivalence should be conducted in both ITT and PP data sets and results are expected to be consistent.
2	**A. Acceptance of MRI parameters as primary outcome measures** EBE recognises that there are differences in the scientific opinion on the value of MRI as a surrogate for clinical outcomes. If MRI parameters are to be accepted as primary outcome measures they need to be performed as accurately and reliably as possible (see comments on line 166-167). In addition the MRI data needs to be supported by clinical outcome data measured over a sufficient time period (see specific comments on line 135-136, 151-154, 158-159 and 164-166). It should be emphasised that different clinical outcomes over 12, 18 and 24 months would raise doubts on any claim of biosimilarity.	Comparable clinical efficacy outcomes up to 12 months are considered sufficient to confirm biosimilarity since these are only part of the data package; quality, non-clinical, PK & PD, safety/immunogenicity data should all support similarity of the biosimilar and reference products. Post-authorisation immunogenicity and impact on efficacy data are only required for an additional 6 months (at least).
2	**B. Immunogenicity / Neutralizing activity** Development of antibodies to interferon beta (IFN-B) is known to vary substantially between the various IFN-B reference medicinal products and patients with respect to frequency and time to	See specific comments.

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	occurrence in the first 2 treatment years and possible reversion of antibody positivity even years beyond this time point (see expert opinion of the NABINMS consortium: Polman et al., Neutralising antibodies to INF-B therapy for multiple sclerosis: recommendations for clinical use, Lancet Neurology, Vol 9 July 2010, 740-750). Moreover, the neutralizing impact of IFN-B induced antibodies on the pharmacodynamic effects and on radiological/clinical outcomes of IFN-B is different for the various IFN-B products (Goodin et al. Neutralizing antibodies to interferon beta: Assessment of their clinical and radiographic impact: An evidence report. Neurology 2007;68:977–984). Therefore, clinical studies for a biosimilar should generate comprehensive data on: A) Frequency, time course and possible reversion of neutralizing activity over a sufficient period of time (see comments on line 186-188) B) The impact of neutralizing activity on pharmacodynamic markers for IFN-B (see comments on line 196-198 and 208-210) C) The potential impact of neutralizing activity on the main outcome parameters (see comments on line 201-207 and 208-210).	
2	**C. Pharmacokinetics** Besides pharmacodynamics (PD), similar pharmacokinetics (PK) of the originator and the biosimilar is the basis for a robust comparability exercise. This applies in particular for IFN-B, as there is	

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	currently no biological response marker identified which is related to the mechanism by which IFN-B influences the clinical evolution of MS, and MRI is not known to be a robust surrogate for clinical outcomes. PK assessment depends primarily on the availability of valid bioassays allowing the determination of the serum concentrations of administered IFN-B over a sufficient period of time. There are principally two approaches to measure IFN-B concentrations in serum: (a) biological IFN activity assays (e.g. MxA gene induction, CPE) and (b) IFN protein mass detection assays (ELISA). In contrast to small molecules, the protein mass is not necessarily correlated with activity for two different products. Hence, similar PK with respect to protein mass does not allow a direct conclusion concerning the pharmacological activity of the compounds. The applicant should justify the rationale for the choice of assay proposed to compare bioavailability of active IFN protein for different products in the proposed PK study. The assays should be sensitive enough to allow quantification of administered IFN-B at least over the entire desired dosing interval.	Agreed
2	**D. Pharmacodynamics** Quantification of pharmacodynamic markers for *in vivo* pharmacological response is methodologically a diverse area with respect to the markers themselves and to quality of analytical methods. To compare the pharmacodynamic (PD) response to administered IFN-B for different products, a major prerequisite is to select markers with a linear dose-response relationship in the therapeutic dose range. This is not the case for all markers often used and mentioned in the guideline. To ameliorate this problem,	

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	multiple response markers should be investigated. Furthermore, it must be ensured that the assay for quantifying the selected markers is robust and delivers reproducible results, i.e. it has to be validated. As MS is a progressive and irreversible disease, without any short-term surrogate marker to demonstrate appropriate efficacy and safety, it is important not to expose MS patients to a product that is not likely to prove biosimilarity. Therefore, after completing the PK/PD aspects of the comparability exercise, a critical assessment to confirm a high likelihood of demonstrating biosimilarity should be undertaken before initiating the comparative clinical efficacy trial, which exposes patients over a long period of time to the proposed biosimilar. Sponsors are encouraged to seek scientific advice to assist in confirming that is appropriate to initiate the comparative clinical efficacy trial.	Agreed While this stepwise approach is highly recommended it is the responsibility of the National Competent Authorities and the Ethic Committees to allow for initiation of the efficacy trial. The guideline only provides recommendation on the data package needed at the time of the application to be considered for approval.
2	**E. Extrapolation of indication** It is well understood that relapsing MS becomes more difficult to treat in the more advanced stages. A 5% higher efficacy on relapses could make the difference between a drug that is effective in EDSS>3.5 and one that is not, even though they may share the same MOA. Additionally, the mechanism of action of IFNs is not well understood and different mechanisms could be involved, given the different patterns of the disease, at the different stages of MS. Therefore, demonstration of efficacy and safety in confirmed RRMS should not allow extrapolation, without supportive data, to non-relapsing forms of MS (e.g. SPMS) or to the earlier stage of clinically isolated syndrome (CIS) prior to definite MS diagnosis.	Whatever the plethoric downstream events triggered by interferon, the common starting point is the binding of interferon to its receptor, which is considered the critical step for extrapolation purposes. Furthermore, extrapolation is based on the entire data package including quality, non-clinical, PK & PD, as well as efficacy/safety in RRMS. If all these aspects support similarity, it is considered that the other MS indications can be granted without additional study. Of note, no IFN-β product has an indication for non-relapsing forms of MS.

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
2	**F. Labelling** Guidance should be given on what sections of the biosimilar product label would be inherited from the reference product, e.g. in terms of safety and immunogenicity. It is expected that all AE from the reference product would be applicable to the biosimilar.	Guidelines on the SmPC are not specific to the product. All ADRs to reference products are indeed included in the SmPC of the biosimilars.
2	**G. Regional harmonisation** Harmonization of guidance across regions is highly desirable and there is a potential opportunity because the FDA is currently generating guidance on biosimilars development.	This is outside the scope of the specific guideline and will be addressed in the overarching guideline.
3	**1. Acceptance of MRI parameters as primary outcome measures** *MRI is widely used as a primary outcome measure in MS phase II clinical trials and as a secondary outcome measure in registration trials. It is particularly useful in phase II trials due to the high rate of subclinical lesion detection, which drives its sensitivity as an outcome measure far beyond relapses (Filippi M, Agosta F. Imaging biomarkers in multiple sclerosis. J Magn Reson Imaging 2010; 31:770–788). Acceptance of MRI parameters as primary endpoints in registration trials would require a solid prediction of clinical effects e.g. on relapses by the effects on MRI.* However, the scientific evidence for this is not very robust and the value of MRI as a surrogate for clinical outcomes has been challenged (*Daumer M, Neuhaus A, Morrissey S, et al. MRI as an outcome in multiple sclerosis clinical trials. Neurology 2009; 72:705–711)*. While there are studies demonstrating the value of MRI outcomes as a surrogate	

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	for clinical outcomes, other studies reported only modest correlations between clinical and MRI outcomes. This dissociation between clinical and MRI outcomes is known for many years and has been referred to as the “clinico-radiological paradox in MS” (Barkhof F. The clinico-radiological paradox in multiple sclerosis revisited. Curr Opin Neurol 2002;15:239–245). Therefore, it cannot be excluded that a product claimed to be biosimilar may appear to yield similar effects on MRI, yet not have the same clinical outcome. While the draft guideline envisages clinical outcome as a secondary endpoint, a primary clinical outcome would be preferred and if MRI parameters are to be accepted as primary outcome measures, it should be emphasised that different clinical outcomes over 12, 18 and 24 months would raise doubts on the biosimilarity claim. Moreover, efforts should be undertaken A) to put MRI findings into perspective with relapse related clinical outcomes obtained according to the guideline (CPMP/EWP/561/98 Rev 1) on clinical investigation of medicinal products for the treatment of multiple sclerosis over a sufficient time period (see specific comments on line 135-136, 151-154, 158-159 and 164-166) and B) to perform the MRI measurements as accurately and reliably as possible (see comments on line 166-167). **2. Immunogenicity / Neutralizing activity** Development of antibodies to IFNB is known to vary substantially	Comparative clinical efficacy data are not required beyond 12 months. Agreed

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	between the various IFN β reference medicinal products and patients with respect to frequency and time to occurrence in the first 2 treatment years and possible reversion of antibody positivity even years beyond this time point (see expert opinion of the NABINMS consortium: Polman et al., Neutralising antibodies to interferon-beta therapy for multiple sclerosis: recommendations for clinical use, Lancet Neurology, Vol 9 July 2010, 740-750). Moreover, the neutralizing impact of IFNB-induced antibodies on the pharmacodynamic effects and on radiological/clinical outcomes of IFNB is different for the various IFNB products (Goodin et al. Neutralizing antibodies to interferon beta: Assessment of their clinical and radiographic impact: An evidence report. Neurology 2007;68:977–984). Therefore, clinical studies for a biosimilar should generate comprehensive data on D) Frequency, time course and possible reversion of neutralizing activity over a sufficient period of time (see comments on line 186-188) and on the impact of neutralizing activity on pharmacodynamic markers for IFNB (see comments on line 196-198 and 208-210) and on the potential impact of neutralizing activity on the main outcome parameters (see comments on line 201-207 and 208-210).	See specific comments.
3	**3. Pharmacokinetics** Besides pharmacodynamics (PD), similar pharmacokinetics (PK) of the originator and the biosimilar is the basis for a robust	

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	comparability exercise. This applies in particular for IFN-ß, as there is currently no biological response marker identified which is related to the mechanism by which IFN-β influences the clinical evolution of MS, and as MRI is known to be not a robust surrogate for clinical outcomes. PK assessment depends primarily on the availability of valid bioassays allowing the determination of the serum concentrations of administered IFN-ß over a sufficient period of time. There are principally two approaches to measure IFN-ß concentrations in serum: (a) biological IFN activity assays (e.g. MxA gene induction, CPE) and (b) IFN protein mass detection assays (ELISA). In contrast to small molecules, the protein mass is not necessarily correlated with activity for two different products. Hence, similar PK with respect to protein mass do not allow a direct conclusion concerning the pharmacological activity of the compounds. To compare bioavailability of active IFN protein for different products in the proposed PK study, it is therefore recommended to analyze plasma samples with both approaches, a cell-based biological activity assay and a protein mass quantifying ELISA. The assays should be sensitive enough to allow quantification of administered IFN-ß at least over the entire desired dosing interval.	The Applicant has to justify the appropriateness of the PK assay but the requirement of two assays is considered excessive. The dosing interval may not be sufficient to evaluate the elimination. Ideally, the full PK profile should be characterised.
3	4. Pharmacodynamics Quantification of pharmacodynamic markers for *in vivo* pharmacological response is methodologically a diverse area with respect to the markers themselves and to quality of analytical	

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	methods. To compare the pharmacodynamic (PD) response to administered IFN-ß for different products, a major prerequisite is to select markers with a linear dose-response relationship in the therapeutic dose range. This is not the case for all markers often used and mentioned in the guideline. To ameliorate this problem, multiple response markers should be investigated. Furthermore, it must be ensured that the assay for quantifying the selected markers is robust and delivers reproducible results, i.e. it has to be validated.	Agreed
3	As MS is a progressive and irreversible disease, without any short-term surrogate marker to demonstrate appropriate efficacy and safety, it is important not to expose MS patients to a product that is not likely to prove biosimilarity. Therefore, after completing the PK/PD aspects of the comparability exercise, a critical assessment to confirm a high likelihood of demonstrating biosimilarity should be undertaken before initiating the comparative clinical efficacy trial, which exposes patients over a long period of time to the proposed biosimilar. Sponsors are encouraged to seek scientific advice to assist in confirming that is appropriate to initiate	While this stepwise approach is highly recommended it is the responsibility of the National Competent Authorities and the Ethics Committees to allow for initiation of the efficacy trial. The guideline only provides recommendation on the data package needed at the time of the application to be considered for approval.
4	This guideline concerns the non-clinical and clinical requirements for biosimilars of interferon beta (IFN-β). Overall this guideline is well-written and in general agreed. With respect to the clinical part the principle is clear. In the context of a biosimilar exercise a similar efficacy/safety as compared to the reference product, not patient benefit per se has to be shown. This allows bridging to the data base of the reference product for which the patient benefit has already been established. Nevertheless the	

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	formulation of the principle as stated in the document may lead to misinterpretations.	Wording has been modified.
	As there is no PD model in multiple sclerosis showing equivalence based on MRI variables is largely agreed. However, it is questioned whether the proposed study designs indeed sufficiently ensure assay sensitivity. From other products it is known that while different doses may well separate from placebo, the doses themselves may not separate from each other. It needs to be discussed whether a different effect of two different doses should not always be demonstrated to ensure assay sensitivity.	Wording has been clarified.
	The principle to accept MRI as endpoint in studies evaluating bio-similarity is clear. The MRI is a sensitive marker of disease activity and when both active agents have the same responsiveness it can reasonably be assumed that both products are equivalent and clinical efficacy can be bridged. However this principle is not followed to the end as still backup of clinical data i.e. relapse rate is required. This is not entirely consistent and may be reconsidered.	The relapse data are supportive only and are needed to assess the impact of immunogenicity.
	The 12 month follow up is needed to validate whether the population included is a sensitive population i.e. whether accumulation of MRI burden otherwise there would be no assay sensitivity. This makes the remark that the most sensitive patient population (line 160-163) should be included superfluous.	In the protocol, inclusion criteria need to be defined with the aim of targeting the most sensitive population.
	Unresolved is the choice of equivalence margin for combined unique active (CUA) MRI lesions. The choice of equivalence margin is based on a combination of statistical and clinical reasoning. However as this is a relative new endpoint that was not used in innovator studies, data are scarce. Second it is unclear on basis of which clinical	

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	reasoning a margin can be defined as it has been rather difficult to relate MRI outcomes to clinical outcome. However, the uncertainty with respect to the choice of equivalence margin might be less important due to the inclusion of the short term placebo arm, two doses i.e. if both active products separate clearly from placebo and the doses as well, assessment of the relative efficacy of reference and test product on CUA lesions is facilitated.	Agreed.

2. Specific comments on text

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
Lines 4-5	2	Comment: The indication (MS) should be mentioned in the title Proposed Change (if any): "Guideline on similar biological medicinal products containing interferon beta for the treatment of multiple sclerosis"	Not accepted. *As a general principle for biosimilar guidelines, the recommendations are product-specific not indication-specific. Extrapolation to other indications is addressed in section 4.4.*
Lines 39-40	4	*Medicinal products containing recombinant IFN-β are currently indicated for patients with relapsing MS or at high risk of developing MS after a single demyelinating event.* Comment: It is noted patients with a single clinical demyelinating event at risk of developing MS (the risk being determined based on the MRI picture) nowadays would be diagnosed as definite RRMS according to the revised McDonald criteria(2010). Proposed change: Therefore it is recommended to delete the last part of this sentence i.e. *Medicinal products containing recombinant IFN-β are currently indicated for patients with relapsing MS ~~or at high risk of developing MS~~*	Partly accepted. *It is preferred to keep a wording that reflects more closely the SmPC wording of the reference products:* *New wording* *"indicated for patients with relapsing MS including those at high risk of developing MS after a single demyelinating event."*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		~~*after a single demyelinating event*~~.	
Line 46	2	Comment: "modest" should not be used since it is a comparative descriptor used without a reference, i.e. modest compared to what? Therefore the statement should be more factual. Proposed Change (if any): "The clinical effects of recombinant IFN–B are shown by decreases in the frequency of exacerbations by approximately…"	Not accepted. *This is a clinical value judgment expressed in qualitative terms that is based on the knowledge of all treatments of MS (see the quote from the MS guideline hereafter). It is further justified by the actual values provided in the text.*
Line 47		Comment: Inconsistency can be expected if population are different, study not powered etc. It is understood that the agency is referring to IFNs as a class, but this is not in line with IFNs which have shown to slow the progression of disability and have a disability claim in the label. This should not be portrayed as "inconsistent". It is important that it is clarified that some IFNs have shown to slow disability and have a disability claim in the label. Proposed Change (if any): "…and inconsistent results on the progression of disability amongst the IFNs available"	Not accepted. *It is meant more generally for inconsistency across trials as well as products as highlighted in the MS guideline:* *"Approved therapies have been shown to favourably modify the short-term evolution of the disease although the benefit is modest, at the cost of significant inconveniences and side effects and it is not known whether the effect is maintained for years. Differences from placebo are not consistent across trials and the sensitivity of the available scales to measure progression of disability as well as other characteristics of clinical trials in this field do not assure the ability to detect clinically relevant differences."*
Lines 50-51	2	Comment: it is not clear if the statement "…asymptomatic liver and white blood cell abnormalities occur more frequently with subcutaneous products…" can be substantiated (it	Accepted. *New wording* *… with the subcutaneous products at the recommended dose*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		seems to blame the difference in AE severity on the route of administration (SC), when this could actually be related to the low dose, low frequency IFN) Proposed change (if any): clarify that this could be explained by difference in posology and not only the route of administration.	*regimens.*
Lines 82-101	4	Comment: In the non-clinical section the view is expressed that generally no *in vivo* studies are needed. We fully support this view.	No action is proposed. The comment is noted.
Lines 83-84	2	Comment: The introduction to the non-clinical studies is already focusing on in vitro studies, this should be more general, leading to the 2 subsequent subsections, the definition of what is required and what is not required. Replace "pharmaco-toxicological"response with "pharmacological" response because this is a more accurate description. Proposed change (if any): "Non-clinical ~~in vitro~~ studies should be performed before initiating clinical development. These studies should be comparative in nature and should be designed to detect differences in the biological activity and in the ~~pharmaco-toxicological~~ pharmacological response …."	Partly accepted. *Wording has been adjusted to more clearly describe the step-wise approach recommended for non-clinical development.*
Lines 97-98	2	Comment: It is stated: If the outcome of the quality evaluation and/or the *in vitro*	Partly accepted.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		bioassays/pharmacological studies raises concerns, the need for additional studies should be considered. It is not explained what “concerns” means in this context. One potential concern is obviously a deviation from "comparability" in the *in vitro* bioassays/pharmacological studies, and it would be useful to state this. Also, it would be appropriate to note that criteria for “comparability” should be pre-defined and adhered to.	*Wording has been adjusted to better explain the meaning of “concerns”.* *However, it may not always be possible to predefine definite comparability criteria on the non-clinical level for each assay applied.*
Lines 97-98	3	It is stated: If the outcome of the quality evaluation and/or the *in vitro* bioassays/pharmacological studies raises concerns, the need for additional studies should be considered. *Comment:* It is not explained what “concerns” means in this context. One potential concern is obviously a deviation from "comparability" in the *in vitro* bioassays/pharmacological studies, and it would be useful to state this. Also, it would be appropriate to note that criteria for “comparability” should be pre-defined and adhered to.	Partly accepted. *Wording has been adjusted to better explain the meaning of “concerns”.* *However, it may not always be possible to predefine definite comparability criteria on the non-clinical level for each assay applied.*
Line 99	3	*Comment*: The suggested strategy to compare the similar biological medicinal product and the reference medicinal product by detailed quality and in vitro characterization (including safety and/or efficacy endpoints) appears appropriate and is in accordance with other guidances focusing on the comparability of	The comment is noted.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		biotechnology derived products. Only in cases where differences are either identified or suspected preclinical in vivo studies should be considered. In such cases it is suggested to specify the species to be used and the biological response markers to be evaluated in the study based on the data generated with the reference product. *Proposed change (addition after Line 99):* In such cases the monkey would be the preferred species and a minimum duration of the study of 4 weeks. Biological response markers, like body temperature, neopterin and Mx protein should be included in such studies in addition to the well-validated toxicological endpoints and additional markers for immunogenicity.	Not accepted. *The decision on an appropriate in vivo programme should be made on a case-by-case basis, depending on the outcome of the quality evaluation and the in vitro bioassays/ pharmacological studies (see also the following comment).*
Lines 99-101	2	Comment: Line 96 clearly states that generally, in vivo studies in animals are not required. Some in vivo studies may be done only if there is concern from the quality/in vitro package. Considering the statement in line 96, we suggest rewording line 99 to 101 to state that any in vivo study has to be adequately justified in relation to the relevant additional information they should provide. Proposed change (if any): "In vivo studies should be designed to specifically address the concern identified. These could include an in vivo pharmacological study and/or a general repeated dose toxicity study in a pharmacologically responsive animal species. In vivo	Accepted. *Wording has been modified accordingly.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		studies not expected to provide relevant additional information should be omitted."	
Lines 100-101	4	Comment: It is expressed that when it can be justified that no additional relevant information is to be gained by performing a study in a pharmacological responsive species, such studies may be omitted. In principle this view is supported, but we think that in view of recent legislation concerning the use of animals for research purposes (Directive 2010/63/EU) this sentence should be rephrased to express that such studies can only be performed when it can be justified by a need for indispensable information that can not be gained otherwise. The wording as it is in the draft leaves the choice to do an in vivo study anyway. The Directive is much stronger in its wording. Proposed change (if any): Replace "*If it can be justified that further studies in a pharmacologically responsive animal species are not expected to provide relevant additional information, then such studies may be omitted*." With "*Further studies in a pharmacologically responsive animal species should only be considered when it is*	Accepted. *Wording has been modified accordingly.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		expected that such studies would provide relevant additional information."	
Lines 105-117	3	*Comment and rationale:* The proposed study design for the pharmacokinetic (PK) study appears appropriate for comparing the PK properties of a biosimilar with the reference product. A single dose approach appears, however, only justified, if (a) the biosimilar and the reference product do not show any accumulation when administered as multiple doses in the desired dosing interval, and (b) it can be proven that both, biosimilar and reference product, result in identical exposure and are eliminated with the same half-life after single dose administration. Due to the limitations of the bioanalytical methods (see below), in particular (b) might be difficult to achieve. If in doubt that multiple dose PK can be indubitably be predicted from single dose data, a multiple dose regimen should be choosen for comparison of the two products. *Proposed change (addition after Line 110):* A single dose approach is only justified, if (a) the biosimilar and the reference product do not show any accumulation after multiple doses in the desired dosing interval, and (b) administration of both, biosimilar or reference product, results in the same exposure and are eliminated with the same half-life. If multiple dose PK cannot be	Partly accepted. *A single-dose cross-over design is the most sensitive design to detect differences as long as the full PK profile can be evaluated.* *New added wording* *The choice of a single or repeated dose (e.g. three doses over a week) regimen should be justified; a single dose is preferred as long as the bioanalytical method is sufficiently sensitive to characterise the full PK profile.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		predicted unambiguously from single dose data, a multiple dose regimen should be chosen for comparison of the two products.	
Lines 106 - 107		Comment: The proposed cross-over design of the PK studies needs to be challenged. Although anti-IFN-B antibodies usually occur only after some months, if they did form earlier they might impair the PK in a cross-over design. A parallel design of the PK study would exclude any potential impact of anti-IFN-B antibodies. Proposed change (if any): Replace "crossover" by "parallel design" and add "Crossover studies are not recommended since antibodies could impair PK results"	Partly accepted. *The advantage of a crossover design outweighs the very low risk of an early immune response, which, if observed, would challenge similarity.* *New added wording* *Although antibody development is not expected after a few doses of IFN-β, their determination should be carried out before/after each treatment course in order to exclude any potential interference with the PK profile.*
Line 108-109	1	Comment: "*The selected dose should be in the sensitive part of the dose-concentration curve.*" **How should the remark "sensitive part" be interpreted?** The sensitive part of the dose concentration curve would normally be the steep part. However, for the currently available IFN-beta's there is limited information on the pharmacokinetics. In addition, serum concentrations of IFN-beta are very low and will be difficult to determine, as indicated in the guideline.	Partly accepted. *The onus is on the Applicant to justify the choice of the dose in the sensitive part of the dose-concentration curve; alternatively, more than one dose can be tested.* *New added wording* *The selected dose should be in the sensitive part of the dose-concentration curve; if available information on the reference product is too scarce, more than one dose should preferably be tested.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
Lines 108 - 109	2	Comment: Clarify that the dose levels used will allow the accurate detection of differences in actual concentration profiles, rather than those which are imputed. It is mentioned that the dose should be selected in the sensitive part of the dose-concentration curve. This is not clear. Do we refer to dose-AUC curve or to dose-response curve? As mentioned later (lines 118-122), serum concentrations are low and the administered dose cannot be increased due to side effects. The best solution is to run the PK study with the therapeutic dose of the reference product. Proposed change (if any): "The selected dose should be the therapeutic dose of the reference product"	Not accepted. *As indicated, dose-concentration curve is meant and not dose-response (PD) curve.* *The selected dose should not necessarily be the therapeutic dose. There is a range of doses that may be administered to healthy volunteers without compromising their safety. The dose should be chosen based on the knowledge of the PKs of the reference product (see previous comment).*
Lines 109 - 110	2	Comment: A single dose approach may only be justified if (a) the biosimilar and the reference product do not show any accumulation when administered as multiple doses in the desired dosing interval, and (b) it can be proven that both, biosimilar and reference product, result in identical exposure and are eliminated with the same half-life after single dose administration. Due to the limitations of the bioanalytical methods (see below), in particular (b) might be difficult to achieve. If in doubt that multiple dose PK can be reliably and accurately predicted from single dose data, a multiple dose regimen should be chosen for	Partly accepted. *A single-dose cross-over design is the most sensitive design to detect differences as long as the full PK profile can be evaluated.* *New added wording* *The choice of a single or repeated dose (e.g. three doses over a week) regimen should be justified; a single dose is preferred as long as the bioanalytical method is sufficiently sensitive to characterise the full PK profile.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		comparison of the two products. *Proposed change (addition after Line 110):* A single dose approach is only justified, if (a) the biosimilar and the reference product do not show any accumulation after multiple doses in the desired dosing interval, and (b) administration of both, biosimilar or reference product, results in the same exposure and are eliminated with the same half-life. If multiple dose PK cannot be predicted reliably and accurately from single dose data, a multiple dose regimen should be chosen for comparison of the two products.	
Line 113	2	Comment: CL/F is irrelevant without a direct measurement of F. We already test the AUC so the same result will be obtained for CL/F. Proposed change (if any): Delete CL/F.	Accepted. *Replaced by clearance.*
Lines 118-122	2	Comment: The current draft guideline suggests two bioanalytical methods for the determination of IFN-B in serum: the CPE assay and ELISA assays. The wording suggests that the ELISA might be superior due to higher sensitivity. Firstly, this might not actually be the case as the lower limit of quantification (LLOQ) in serum samples to measure exogenous IFN-B is much higher (~60 pg/mL) than the value in buffer (~5 pg/mL) which is advertised by the assay providers. This is due to endogenous IFN and matrix effects of the serum. Furthermore, the principal difference of	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		analysis should be clarified in the guideline: The CPE assay determines the biological activity of IFN-B present in the serum sample; whereas the ELISA assays determine the mass of IFN protein in a serum sample. With respect to efficacy, the activity assay is considered more relevant. However, to compare bioavailability of active IFN-B protein, and due to the potential disconnect between biological activity and protein mass (see General Comments) the applicant should justify the rationale for the choice of assay.	Accepted. *This wording has been added.*
		With respect to activity assays, the CPE assay can currently not be recommended for a comparative PK study using a bioequivalence approach due to its high LLOQ (resulting in insufficient period covered for PK comparison), low and variable accuracy (75-130%) and high inter-assay variability (as high as about 30%). The cell based MxA gene induction assay in contrast, which determines similar to the CPE assay the IFN-B activity in serum samples when added to the cell culture, has higher precision and accuracy. The MxA assay should therefore be recommended or at least mentioned.	Accepted.
		In any case, the bioanalytical method sensitivity should allow quantifying INF-ß concentrations at least over the entire dosing interval, e.g. 2 days for Betaferon.	
		Proposed change (additions in Lines 120-122): … that determination is possible with a cell based myxovirus	Accepted.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		resistance protein A (MxA) induction bioassay or the cytopathic effect (CPE) bioassay, which both determine the biological activity of the administered IFN-ß in serum samples. Due to higher precision and accuracy, the MxA induction assay should be preferred. Recently, more sensitive ELISA assays, which determine the IFN-ß protein mass, have been developed that allow determination of concentration as low as the pg level per mL. Therefore, the applicant should justify the rationale for the choice of assay to quantify the pharmacokinetics of the protein as well as its activity. All assays used should allow quantifying INF-ß concentrations at least over the entire dosing interval.	*The CPE assay has been removed.* *New wording* *Possible methods of detection include a cell-based myxovirus resistance protein A (MxA) induction assay, which measures the biological activity of IFN-β in serum samples, and ELISA assays, which determine the IFN-β protein mass. The Applicant should justify the rationale for the choice of assay.* Partly accepted. *The requirement for two assays (activity and mass) is considered excessive.* *The dosing interval may not be sufficient to evaluate the elimination. Ideally, the full PK profile including elimination should be characterised.*
Lines 121-122	3	*Comment and rationale:* The current draft guideline suggests two bioanalytical methods for the determination of IFN-ß in serum: the CPE assay and ELISA assays. The wording suggests that the ELISA might be superior due to higher sensitivity. Firstly, this might not actually be the case as the lower limit of quantification (LLOQ) in serum samples to measure exogenous IFN-ß is much higher (~60 pg/mL) than the value in buffer (~5 pg/mL) which is advertised by the assay providers. This is due to endogenous IFN and matrix effects of the serum. Furthermore, the principal	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		difference of analysis should be clarified in the guideline: The CPE assay determines the biological activity of IFN-ß present in the serum sample. ELISA assays determine the mass of IFN protein in a serum sample. With respect to efficacy, the activity assay is considered more relevant. However, to compare bioavailability of active IFN beta protein, and due to the potential disconnect between biological activity and protein mass (see General Comments) it would be necessary to analyze plasma samples with both approaches, an activity assay and a protein mass quantifying ELISA, in the proposed PK study. With respect to activity assays, the CPE assay can currently not be recommended for a comparative PK study using a bioequivalence approach due to its high LLOQ (resulting in insufficient period covered for PK comparison), low and variable accuracy (75-130%) and high inter-assay variability (as high as about 30%). The cell based MxA gene induction assay in contrast, which determines similar to the CPE assay the IFN-ß activity in serum samples when added to the cell culture, has higher precision and accuracy. The MxA assay should therefore be recommended or at least mentioned. In any case, the bioanalytical method sensitivity should allow quantifying INF-ß concentrations at least over the entire dosing interval, e.g. 2 days for Betaferon.	Not accepted. *The requirement for two assays (activity and mass) is considered excessive.* Accepted. *The CPE assay has been removed.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (additions in Lines 120-122): … that determination is possible with a cell based myxovirus resistance protein A (MxA) induction bioassay or the cytopathic effect (CPE) bioassay, which both determine the biological activity of the administered IFN-ß in serum samples Due to higher precision and accuracy, the MxA induction assay should be preferred. Recently, more sensitive ELISA assays, which determine the IFN-ß protein mass, have been developed that allow determination of concentration as low as the pg level per mL. This should be used in addition to the activity assay, to quantify the pharmacokinetics of the protein as well as its activity. All assays used should allow quantifying INF-ß concentrations at least over the entire dosing interval.	Partly accepted. *New wording* *Possible methods of detection include a cell-based myxovirus resistance protein A (MxA) induction assay, which measures the biological activity of IFN-β in serum samples, and ELISA assays, which determine the IFN-β protein mass. The Applicant should justify the rationale for the choice of assay.* *The dosing interval may not be sufficient to evaluate the elimination. Ideally, the full PK profile should be characterised.*
Line 124	2	*Comment:* Validated assays should be used for quantification of pharmacodynamic markers. *Proposed change*: …be evaluated using validated assays as part of the … Comment: It is unclear if there is an expectation to demonstrate equivalence of the pharmacodynamic (PD) markers. It is recommended that the PD parameters should be equivalent; if not, this should raise the same concerns as with differences in neutralising antibodies (NAbs). Methods for comparing the PD properties of the biosimilar and RP, and the	Accepted. Not accepted. *Formal equivalence testing of PD parameters is not considered feasible (especially due to the difficulties of justifying an equivalence margin) or needed since the demonstration of equivalent efficacy relies on the pivotal trial in MS patients. Only descriptive comparisons and a discussion*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		equivalence margins should all be defined *a priori* and justified.	*of the observed differences are expected.*
Line 124	3	*Comment:* Validated assays should be used for quantification of pharmacodynamic markers. *Proposed change*: ...be evaluated using validated assays as part of the …	Accepted.
Lines 126 - 133	2	*Comment:* Some of the pharmacodynamic (PD) markers do not have a linear dose-response relationship around the therapeutic dose range of IFN-ß-1b (e.g. ß2-microglobulin,….). Consequently, different doses of IFN-ß-1b can induce similar pharmacodynamic effects. Such markers are not suitable for a reliable comparison of the PD effect of different products. The use of multiple PD markers might reduce this problem of lacking dose-response linearity but does not solve it. The preference for the PD MxA induction assay is supported, and this should definitely be listed more emphatically as a marker to be investigated. The most reliable and consistent dose-response relationship was found for neopterin which should, therefore, also be added to this mandatory list. *Proposed change (addition in Line 128):* … ("fingerprint approach"). It should be noted that some of these markers do not exhibit a linear dose-response relationship in the therapeutic dose range of IFN-ß and	Not accepted. *The PD investigation may not be restricted to the therapeutic dose range.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		are therefore not suitable for a comparative evaluation. To ameliorate this problem, multiple response markers should be investigated. Amongst others, ~~these such~~ biological response markers include… *Proposed change (addition after Line 133):* … MxA… should be one of the selected markers. Neopterin was found to show a consistent and robust dose-response relationship and should therefore also be investigated.	*The fingerprint approach is already recommended.* Accepted
Line 126-133	3	*Comment:* Some of the pharmacodynamic (PD) markers do not have a linear dose-response relationship around the therapeutic dose range of IFN-ß-1b (e.g. ß2-microglobulin,….). Consequently, different doses of IFN-ß-1b can induce similar pharmacodynamic effects. Such markers are not suitable for a reliable comparison of the PD effect of different products. The use of multiple PD markers might reduce this problem of lacking dose-response linearity but does not solve it. The preference for the PD MxA induction assay is supported, and this should definitely be listed more emphatically as a marker to be investigated. The most reliable and consistent does-response relationship was found for neopterin which should, therefore, also be added to this mandatory list. . *Proposed change (addition in Line 128):* … ("fingerprint	Not accepted.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		approach"). It should be noted that some of these markers do not exhibit a linear dose-response relationship in the therapeutic dose range of IFN-ß and are therefore not suitable for a comparative evaluation. To ameliorate this problem, multiple response markers should be investigated. Amongst others, ~~these such~~ biological response markers include… *Proposed change (addition after Line 133):* … MxA… should be one of the selected markers. Neopterin was found to show a consistent and robust dose-response relationship and should therefore also be investigated.	*The PD investigation may not be restricted to the therapeutic dose range.* *The fingerprint approach is already recommended.* Accepted
Lines 135-136	2	Comment: Considering that the effect of a biosimilar on relapses will be an important secondary outcome measure needed to support a primary MRI endpoint (see general comment 1A) and that the evaluation of relapses is subject to bias in unblinded study (see CPMP/EWP/561/98 Rev 1), a double-blind design is strongly recommended. If double blinding is technically not possible, which should be justified, alternative measures should be applied to avoid bias. Proposed change: Similar clinical efficacy between the biosimilar and reference medicinal product should be demonstrated in	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		an adequately powered, randomised, parallel group equivalence clinical trial, preferably double-blind. If double blinding is technically not possible, which should be justified, alternative measures should be applied to avoid bias.	Accepted. *New added wording* *If blinding is technically not feasible, alternative measures should be applied to avoid information bias.*
Lines 135-136	3	Comment: Considering that the effect of a biosimilar on relapses will be an important secondary outcome measure needed to support a primary MRI endpoint (see general comment 1A) and that the evaluation of relapses is subject to bias in unblinded study (see CPMP/EWP/561/98 Rev 1), a double-blind design is strongly recommended. If double blinding is technically not possible, which should be justified, alternative measures should be applied to avoid bias. Proposed change: Similar clinical efficacy between the biosimilar and reference medicinal product should be demonstrated in an adequately powered, randomised, parallel group equivalence clinical trial, preferably double-blind. If double blinding is technically not possible, which should be justified, alternative measures should be applied to avoid bias.	Accepted.
Line 143-145	4	*….since the focus of the biosimilarity exercise is to demonstrate similar efficacy and safety compared to*	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		the reference product, not patient benefit per se, which has already been established by the reference product. It is recommended to rephrase this as the statement *not patient benefit per se* may be taken out of the context and lead to misinterpretation. Proposed change: *… since the focus of the biosimilarity exercise is to demonstrate similar biological properties of the new product to that of the reference product. This allows bridging to the benefit-risk of the reference product.*	Partly accepted. *It is not understood why this comment is misleading. A similar statement was used in other biosimilar guidelines. However, the statement has been slightly reworded.* *New wording* *since the focus of this trial is to demonstrate comparable clinical activity of the biosimilar product to the reference medicinal product, which then allows bridging to the benefit-risk of the reference product.*
Lines 149-154	1	Comment: *"Regarding the study design, assay sensitivity could be shown by a three-arm trial including a placebo arm for a short period of time (e.g. 4 months) sufficient to demonstrate superiority of both the biosimilar and reference products over placebo using an MRI endpoint."* **Is it also possible for the assay sensitivity analysis to combine the biosimilar and reference data and compare this to placebo?** The formulation in the draft guidance suggests that both the biosimilar and the reference product have to be tested individually vs. placebo. Because the aim of the superiority analysis is to show assay sensitivity,	Not accepted. *In the 3-arm trial it is the comparison of the reference to placebo that really shows the assay sensitivity. If the reference product is not better than placebo, it would be difficult to understand what it means to be equivalent to it. Likewise, the biosimilar product needs to be better than placebo or there would be real concerns about the efficacy of*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		rather than to show efficacy of either of the treatments, evaluation of the difference between the average of the active treatments and the placebo group could be sufficient; more precisely, by evaluating the contrast: **(log(mean of TEST group)+log(mean of REFERENCE group))/2 – log(placebo).**	*the biosimilar product. Thus, both biosimilar and reference products need to show superiority to placebo.*
Line 149-157	4	Regarding the study design, it is stated that assay sensitivity could be shown either by a three-arm trial including a placebo arm for a short period of time (e.g. 4 months) or a three-arm trial with the reference product and two doses of the biosimilar product, for which it can be reasonably assumed that they will exhibit differences in MRI and clinical outcomes over time. Comment: It is questioned whether assay sensitivity is sufficiently covered by the proposed three-arm trials. From other products it is known that while different doses may well separate from placebo, the doses themselves may not separate from each other. It needs to be discussed whether a different effect of two different doses should not always be demonstrated to ensure assay sensitivity.	Partly accepted. *Based on the information available for the reference products, it is considered that formal demonstration of superiority of one dose of the reference product to another is unlikely to be achievable in many cases as a subtherapeutic dose cannot be administered for one year. However, a difference in MRI and clinical outcomes is expected to be observed over time to support the assay sensitivity of the trial. The statement has been reworded to clarify that this is "expected to be observed" rather than "assumed".* *<u>New wording</u>* *…for which differences in MRI and clinical outcomes are expected to be observed over 12 months in order to support the assay sensitivity of the trial.*
Lines 151 - 154	2	Comment: Considering that the effect of a biosimilar on MRI findings should be backed-up by clinical	Not accepted. *Placebo treatment for more than 4 months is not considered*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		findings (see general comment 1A), the placebo period for demonstration of superiority of the biosimilar and the reference product should be long enough to detect effect on relapses. Proposed change: Regarding the study design, assay sensitivity could be shown by a three-arm trial including a placebo arm for a ~~short~~ period of time (~~4~~ >6 months) sufficient to demonstrate superiority of both the biosimilar and reference products over placebo using an MRI endpoint supported by a trend in superiority on clinical endpoints.	*acceptable in such patient population as it might seriously impair their chances of improvement under delayed active treatment. It is considered sufficient that assay sensitivity is shown using the primary endpoint (MRI), superiority over placebo being achievable over a 4-month period.*
Lines 151-154	3	Comment: Considering that the effect of a biosimilar on MRI findings should be backed-up by clinical findings (see general comment 1A), the placebo period for demonstration of superiority of the biosimilar and the reference product should be long enough to detect effect on relapses. Proposed change: Regarding the study design, assay sensitivity could be shown by a three-arm trial including a placebo arm for a period of time (>6 months) sufficient to demonstrate superiority of both the biosimilar and reference products over placebo using an MRI endpoint supported by a trend in	Not accepted. *Placebo treatment of more than 4 months is not considered acceptable in such patient population as it might seriously impair their chances of improvement under delayed active treatment. It is considered sufficient that assay sensitivity is shown using the primary endpoint (MRI), superiority over*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		superiority on clinical endpoints.	*placebo being achievable over a 4-month period.*
Line 154	2	Comment: avoid the word crossed over Proposed change (if any): replace “crossed over” by “randomised to an active arm”	Partly accepted. *It is not proposed to randomise placebo patients to either biosimilar or reference product but to switch them all to the biosimilar product to increase the size of the safety database.* *<u>New wording</u>:* *Patients in the placebo arm could be subsequently switched to the biosimilar product …*
Line 157	2	Comment: Be more specific than the words “over time”, indicate 12 months to be consistent with the duration of the trial in line 159. Proposed change (if any): Replace “over time” with “over 12 months”	Accepted. *<u>New wording</u>* *…for which differences in MRI and clinical outcomes are expected to be observed over 12 months in order to support assay sensitivity.*
Lines 158-159	1	Comment: “*Whatever the design, the duration of the trial should be sufficient to show comparable efficacy on MRI end points and provide relevant information on clinical outcomes, i.e. not less than 12 months*” **It is suggested to remove the minimum requirement of 12 months.** If with the available data for the reference product a sponsor can evaluate and calculate the margins, sample size for several durations of the trial, and there	Not accepted. *Based on clinical relapse data available for the reference products, it is considered important to recommend a sufficient duration of follow-up (at least 12 months) to exclude the absence of substantial difference in relapses between the biosimilar and reference product. Of note, statistical power is not required for clinical relapses. Furthermore, 12-month comparative data are also required for immunogenicity.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		is sufficient data available from the reference product, for example at 6 months, to design an appropriate trial this should be possible. Although a minimum duration of 12 months could be preferred it should not a priori be ruled out that other designs with a shorter duration can provide the same level of assurance. Currently the MRI data available for the IFNs will probably lead to a trial with a minimal duration of 12 months, however new data can become available later. Proposed change (if any): *Whatever the design, the duration of the trial should be sufficient to show comparable efficacy on MRI end points and provide relevant information on clinical outcomes~~, i.e. not less than 12 months~~.*	
Lines 158-159	2	Comment: It should be noted that the 12-month data should be available at the time of submission of the MAA, but that a study of longer duration (24 months) is preferred to ensure that appropriately robust, supportive information on relapse rates can be confirmed, post-approval. Proposed change: Whatever the design, the duration of the trial should be sufficient to show comparable efficacy on MRI	Not accepted. *Comparative efficacy data are not considered necessary beyond 12 months provided that quality, non-clinical, PK & PD, safety/immunogenicity data all support similarity of the biosimilar and reference products.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		endpoints and provide relevant information on clinical outcomes, i.e. not less than 12 months available at the time of submission, 18-month data at day 121 responses and 24-month data post-marketing.	
Lines 158-159	3	Comment: It should be noted that the 12-month data should be available at the time of submission of the MAA, but that a study of longer duration (24 months) is preferred to ensure that appropriately robust, supportive information on relapse rates can be confirmed, post-approval. Proposed change: Whatever the design, the duration of the trial should be sufficient to show comparable efficacy on MRI endpoints and provide relevant information on clinical outcomes, i.e. not less than 12 months available at the time of submission and 18-month data on clinical outcomes to be submitted at day 121 as well as 24-month clinical data pre-marketing.	Not accepted. *Comparative efficacy data are not considered necessary beyond 12 months provided that quality, non-clinical, PK & PD, safety/immunogenicity data all support similarity of the biosimilar and reference products*.
Lines 160 - 161	2	Comment: The most sensitive population is the population that is in the label of the reference product or the population of the clinical trial of the reference product. Additionally, The population must be RRMS with a sufficient disease activity based on relapse frequency	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		and / or MRI criteria to anticipate rapid changes in MRI. The literature confirmed that in clinical trials with INF-B, MRI predict at best 80% of the variance for relapses (Sormani MP et al Ann Neurol. 2009 Mar;65(3):268-75). Proposed change (if any): Add "Since the primary efficacy endpoint is MRI, the population should be selected mainly on the MRI activity and not clinical outcomes." and reword: "Additionally, the most sensitive population, as described in the label, or, in the pivotal clinical trials of the reference product, which would enable the detection of differences between the biosimilar and reference products, should be selected."	Not accepted. *The onus is on the Applicant to justify the choice of the most sensitive clinical model, which does not necessarily correspond to the exact label or original pivotal studies of the reference product.* *The current wording for the patient selection criteria refers to "sufficient disease activity based on relapse frequency and/or MRI criteria to anticipate rapid changes in MRI". It is considered that MRI is sufficiently emphasised in these criteria.*
Lines 161-163	1	Comment: For patient selection it is described that patients should have sufficient disease activity based on relapse frequency and/or MRI criteria to anticipate rapid changes in MRI. **Can the requirements for the inclusion criteria be more elaborated with examples?** For instance the (minimal/maximal) number of T1 Gadolinium enhancing lesions on screening and/or the (minimal/maximal) number of T2 lesions on screening and/or number of relapses in the last (1 or 2 or 3) year(s).	Not accepted. *It is not considered necessary to be more specific in a guideline, especially in such rapidly evolving field as MRI. The onus is on the Applicant to justify the inclusion criteria for the definition of the most sensitive patient population.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
Lines 164-166	2	Comment: Clinical information, in particular, on relapses should be obtained as described in CPMP/EWP/561/98 Rev 1 (see general comment 1A). Recommend the language on clinical outcomes showing the same trend as the MRI variables is more specific i.e. does this refer to a trend in similarity, or, a trend in effect? Proposed change: MRI-based variables are acceptable primary endpoints in the context of a biosimilar comparison if backed up by relapse-related clinical outcomes; no formal equivalence test is required for clinical outcomes, which would be expected to show the same trend in effect as the MRI-based variables. Relapses should be differentiated from pseudo-extrapolation and should be defined accurately as described in the guideline on clinical investigation of medicinal products for the treatment of MS (CPMP/EWP/561/98 Rev 1). Repeated MRI scans should be…	Accepted. *New wording:* *MRI-based variables are acceptable primary endpoints in the context of a biosimilar comparison if backed up by relapse-related clinical outcomes; no formal equivalence test is required for clinical outcomes, which would be expected to show the same trend in effect as the MRI-based variables. A relapse should be differentiated from a pseudo-exacerbation and accurately defined.*
Lines 164-166	3	Comment: Clinical information, in particular, on relapses should be obtained as described in CPMP/EWP/561/98 Rev 1 (see general comment 1A).	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change: MRI-based variables are acceptable primary endpoints in the context of a biosimilar comparison if backed up by relapse-related clinical outcomes; no formal equivalence test is required for clinical outcomes, which would be expected to show the same trend as the MRI-based variables. Relapses should be differentiated from pseudo-extrapolation and should be defined accurately as described in the guideline on clinical investigation of medicinal products for the treatment of MS (CPMP/EWP/561/98 Rev 1). Repeated MRI scans should be…	Accepted. *New wording:* *MRI-based variables are acceptable primary endpoints in the context of a biosimilar comparison if backed up by relapse-related clinical outcomes; no formal equivalence test is required for clinical outcomes, which would be expected to show the same trend in effect as the MRI-based variables. A relapse should be differentiated from a pseudo-exacerbation and accurately defined.*
Lines 166-167	2	Comment: It should be specified what "repeated MRI" means. The frequency of the MRI should be the same as measured in the pivotal studies of the reference product or every 3 months. Considering the importance of MRI-based variables, MRI measurements should be performed as accurately and reliably as possible (see the general comment 1B and CPMP/EWP/561/98 Rev 1). Proposed change: Repeated MRI scans, should be performed during the trial as measured in the pivotal studies of the reference product or every 3 months. All possible actions should be taken to ensure high quality MRI data and maximum reliability of measurements.	Partly accepted. *It is not considered that the guideline needs to be more specific on the frequency of MRI examination. The onus is on the Applicant to define this frequency according to the selected design and endpoints.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Updated recommendations on appropriate technical facilities and standardized procedures and training should be followed. The reading of the images should be central and blinded.	*The statement on the quality of the MRI data has been added.*
Lines 166-167	3	Comment: Considering the importance of MRI-based variables, MRI measurements should be performed as accurately and reliably as possible (see the general comment 1B and CPMP/EWP/561/98 Rev 1). Proposed change: Repeated MRI scans should be performed during the trial. All possible actions should be taken to ensure high quality MRI data and maximum reliability of measurements. Updated recommendations on appropriate technical facilities and standardized procedures and training should be followed. The reading of the images should be central and blinded.	Accepted. *The statement on the quality of the MRI data has been added.*
Line 168		Comment: We agree that CUA makes sense because it is a good combination of the lesion load and the inflammatory component. It is not easy to standardize them and mainly Rebif has the CUA. The other IFNs refer to Gd-enhancing lesions and/or new or active T2 lesions. Proposed change (if any): "The most sensitive documented MRI variable ~~is the~~	Partly accepted. *If justified, the use of a slightly different endpoint compared to the reference product's historical data is acceptable (the most sensitive endpoint has been chosen based on MRI expert recommendations) but the wording has been changed to be more flexible regarding other MRI variables.* *New wording*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		~~combined unique active lesions (CUA, defined as new….without double counting)~~; would be those used in the pivotal studies of the reference product; a cumulative estimate……"	*The combined unique active lesions (CUA, defined as new gadolinium-enhancing T1-weighted lesions and new/enlarging T2-weighted lesions without double counting) are the most sensitive documented MRI variable, and therefore, should always be determined; a cumulative estimate over several scans may be calculated. Other MRI variables may also be used as primary endpoint if adequately justified.*
Line 173	2	Comment: The efficacy of the Reference Product relative to placebo should be specifically addressed when justifying the adequacy of the proposed equivalence margins for the biosimilar.	Accepted. *New wording:* *The equivalence margin for the primary MRI endpoint should be pre-specified and adequately justified based on MRI data for the reference medicinal product relative to placebo*
Line 174	2	Comment: If the study lasts 12 months, we do not have high drop out. Proposed change (if any): "…to the potential~~ly high~~ drop-out rate and the way of handling missing data."	Accepted.
Lines 177-178	2	Comment: Clinical trial data are sufficient to detect the most common adverse drug reactions for a drug, uncommon or rare adverse drug reactions usually occur during postmarketing. This information should be included here. Proposed change: ~~Comparative safety data from the efficacy trial are usually sufficient to provide an adequate premarketing safety database, and therefore~~	Accepted.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		~~should allow for reassurance of safety prior to marketing authorisation.~~ Comparative safety data from the efficacy trial are usually **only** sufficient to investigate the more frequent adverse events data, and these provide an adequate pre-approval safety database for such events, but not for rarer adverse effects	*New wording:* *Comparative safety data from the efficacy trial are usually sufficient to investigate the more frequent adverse reactions and provide an adequate pre-authorisation safety database for such reactions but not for rarer adverse reactions, which should be addressed post-authorisation.*
Lines 177 to 178	3	Comment: Clinical trial data are sufficient to detect the most common adverse drug reactions for a drug, uncommon or rare adverse drug reactions usually occur during postmarketing. This information should be included here. Proposed change: ~~Comparative safety data from the efficacy trial are usually sufficient to provide an adequate premarketing safety database, and therefore should allow for reassurance of safety prior to marketing authorisation.~~ Comparative safety data from the efficacy trial are usually **only** sufficient to investigate the more frequent adverse events data, and these provide an adequate pre-approval safety database for such events, but not for rarer adverse effects.	Accepted. *New wording:* *Comparative safety data from the efficacy trial are usually sufficient to investigate the more frequent adverse reactions and provide an adequate pre-authorisation safety database for such reactions but not for rarer adverse reactions, which should be addressed post-authorisation.*
Lines 179-180	2	Comment: Influenza-like symptoms or injection site reactions etc. are the most frequent ADRs, but not AEs of special interest. Also the most important laboratory	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		test abnormalities should be specified. Proposed change: ~~Adverse events of specific interest include influenza-like symptoms, injection reactions and laboratory test abnormalities.~~ The most frequently observed adverse reactions are a flu-like symptom complex (fever, chills, arthralgia, malaise, sweating, headache, or myalgia), injection site reactions or laboratory test abnormalities, including complete blood count, transaminases and thyroid function tests. In addition adverse events of special interests include hepatotoxicity and depression, which should be included in the risk management plan.	Partly accepted. *The most frequent ADRs are already listed in section 1.* *The AEs of special interests are now mentioned in the pharmacovigilance plan (section 4.3).*
Lines 179 to 180	3	Comment: Influenza-like symptoms or injection site reactions ect. are the most frequent ADRs, but not AEs of special interest. Proposed change: ~~Adverse events of specific interest include influenza-like symptoms, injection reactions and laboratory test abnormalities.~~ The most frequently observed adverse reactions are a flu-like symptom complex (fever, chills, arthralgia, malaise, sweating, headache, or myalgia), injection site reactions or laboratory test abnormalities. In addition adverse events of special interests include	Partly accepted. *The most frequent ADRs are already listed in section 1.* *The AEs of special interests are now mentioned in the pharmacovigilance plan (section 4.3).*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		hepatotoxicity and depression, which should be included in the risk management plan.	
Lines 186-191	2	Comment: NAbs should be followed-up for 18 months. The strategy to measure the NAbs at regular intervals should be aligned with the data available. NAbs positive start to be available at month 3 until month 18. There is no need to assess NAbs every month at the beginning. Assessments at month 3 and then every 6 months for a total of 24 months should be sufficient because, to detect positive NAbs at 2 consecutive visits, 3 months apart, will require up to 24 month follow-up. Proposed change: A minimum of 12 month comparative immunogenicity data should be submitted pre-authorisation with further assessment to be continued ~~post-approval for at least 6 months~~ for the biosimilar product with 18-month data to be submitted at day 121 and 24-month data after approval. A strategy that includes serum sampling at baseline and at regular intervals is necessary for assessing the comparability of the dynamics of antibody development during therapy, e.g., at month 3 and then every 6 months. To detect positive NAb at 2 consecutive visits, 3 months apart, will require up to 24 month follow-up."	Not accepted. *It is considered that a total of 18-month immunogenicity follow-up is sufficient provided the whole data package including quality, non-clinical, PK & PD, efficacy/safety, and immunogenicity data up to 12 months (pre-authorisation), all support similarity of the biosimilar and reference products.* *The guideline only provides an example of possible sampling schedule; the onus is on the Applicant to justify the selected strategy. It is considered that the rate of antibody development at the beginning of the treatment (first 3 months) is adding sensitivity to the comparison of the developing immune response.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
Lines 186-188	3	Comment: The development and time course of IFNB-induced antibodies varies considerably in the first 2 treatment years as outlined in our second general comment. Therefore, a minimum period of 24 months is recommended for monitoring of immunogenicity including neutralizing antibodies in patients of the clinical studies. Proposed change: A minimum of 12 month comparative immunogenicity data should be submitted pre-authorisation with further assessment to be continued for the biosimilar product with 18-month data to be submitted at day 121 and 24-month data, pre marketing..	Not accepted. *It is considered that a total of 18-month immunogenicity follow-up is sufficient provided the whole data package including quality, non-clinical, PK & PD, efficacy/safety, and immunogenicity data up to 12 months (pre-authorisation), all support similarity of the biosimilar and reference products*.
Lines 193-195	2	Comment: Capturing IFN-ß using a monoclonal antibody would also mask some potential binding epitopes of anti-drug antibodies. Proposed change (if any): Delete sentence “e.g., ELISAs using a monoclonal antibody to capture IFN-β”	Accepted.
Lines 196-198	2	Comment: The impact of neutralizing activity on the pharmacodynamics of the biosimilar should be quantified in patients of the clinical studies. In order to fully understand any differences in	Not accepted. *Although a valuable tool to assess the PD impact of neutralising antibodies, systematic measurement of MxA expression is considered excessive in this context as it would*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		immunogenicity in the clinic, it is recommended that PD assessments should be correlated with NAb results. PD assessments may be more sensitive to detecting presence of NAbs and therefore should be analysed carefully to detect absence of difference in PD marker up regulation between the biosimilar and RP. Proposed change: It is recommended that the standardised MxA protein NAb assay or a NAb assay that has been validated against the MxA protein NAb assay is used (EMEA/CHMP/BWP/580136/2007). In patients of the clinical studies with neutralizing activity, the pharmacodynamic effect of the biosimilar should be quantified, e.g. by measuring MxA-expression.	*require measurement at baseline and at an early endpoint (to quantify the activity of the interferon) in all patients to be able to interpret the results in patients developing neutralising antibodies. Moreover, their impact on MRI and clinical relapses is evaluated.*
Lines 196-198	2	Comment: The impact of neutralizing activity on the pharmacodynamics of the biosimilar should be quantified in patients of the clinical studies (see general comment 2B). Proposed change: It is recommended that the standardised MxA protein NAb assay or a NAb assay that has been validated against the MxA protein NAb assay is used (EMEA/CHMP/BWP/580136/2007). In patients of the clinical studies with neutralizing activity, the pharmacodynamic effect of the biosimilar should be	Not accepted. *Although a valuable tool to assess the PD impact of neutralising antibodies, systematic measurement of MxA expression is considered excessive in this context as it would require measurement at baseline and at an early endpoint (to quantify the activity of the interferon) in all patients to be able to interpret the results in patients developing neutralising antibodies. Moreover, their impact on MRI and clinical relapses is evaluated.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		quantified, e.g. by measuring MxA-expression.	
Lines 201-207	2	Comment: Neutralizing activity might impact on the clinical and radiological outcomes under IFN-B therapy. Therefore, it is recommended to compare the main outcomes of the clinical study in NAb positive versus NAb negative patients. Samples negative for anti-IFN-ß antibodies should be identified based on a cut-point not based on dilutions or titres. Proposed change: Finally, patients should be categorised according to the evolution of their immune response over time using predefined criteria and the primary/secondary outcomes of the clinical study should be compared between these categories. For example, the patient's Nab status may be defined as antibody negative (-ve for all post-treatment samples according to predefined ~~low/high dilutions or titres~~ cut off point) or antibody positive, which can be categorised as 'transiently positive' (1 or more post-treatment samples +ve, followed by –ve samples at all subsequent and at least 2 sampling time points) or 'persistently positive' (2 or more consecutive post-treatment samples consistently +ve).	Partly accepted. *New wording:* *MRI activity and clinical relapses should be compared between these categories for both the biosimilar and reference product. The impact of NAbs on clinical outcomes is unlikely to be sufficiently ascertainable before 12 months of therapy and thus will need to be further evaluated post-authorisation as part of the risk management plan.* *This has not been changed as it is meant to emphasise the need for several cut-off points to define low and high titres.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
Lines 201-207	3	Comment: Neutralizing activity might impact on the clinical and radiological outcomes under IFNB therapy (see general comment 2C). Therefore, it is recommended to compare the main outcomes of the clinical study in NAb positive versus NAb negative patients. Proposed change: Finally, patients should be categorised according to the evolution of their immune response over time using predefined criteria and the primary/secondary outcomes of the clinical study should be compared between these categories. For example, the patient's Nab status may be defined as antibody negative (-ve for all post-treatment samples according to predefined low/high dilutions or titres) or antibody positive, which can be categorised as 'transiently positive' (1 or more post-treatment samples +ve, followed by –ve samples at all subsequent and at least 2 sampling time points) or 'persistently positive' (2 or more consecutive post-treatment samples consistently +ve).	Accepted. *<u>New wording:</u>* *MRI activity and clinical relapses should be compared between these categories for both the biosimilar and reference product. The impact of NAbs on clinical outcomes is unlikely to be sufficiently ascertainable before 12 months of therapy and thus will need to be further evaluated post-authorisation as part of the risk management plan.*
Lines 208-210	2	Comment: The concept of biosimilarity would also be contradicted, if neutralizing antibodies would have a substantial effect on the pharmacodynamic effect of IFN-B or on the outcome measures of the clinical	Partly accepted. *It is known that development of NAbs has an effect on treatment efficacy, thus an impact of NAbs on PD/efficacy is indeed expected. The concept of biosimilarity would only be contradicted if this impact differs between the biosimilar and*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		study. In addition, the frequency of Nabs should be comparable since differences would be indicative of dissimilar quality attributes that may impact potency and safety. Proposed change: Although the clinical impact of binding, non-neutralising antibodies is not clear, an increased or decreased frequency of such antibodies for the test product relative to the reference product would be indicative of significant product dissimilarity and contradict the concept of biosimilarity. Moreover, it would be of concern, if neutralizing antibodies have a substantial impact on the pharmacodynamic effect of the biosimilar or on the outcomes of the clinical study.	*reference product.* *A decreased immunogenicity would be acceptable and would not preclude biosimilarity provided this is the only difference observed (i.e. similar quality, non-clinical, PK & PD, efficacy/safety). A clarification statement has been added.* *New wording* *The immune response to the biosimilar and reference medicinal products is expected to be comparable with regard to the incidence and titres of antibodies (neutralising or not) as well as their impact on efficacy; although the clinical impact of binding, non-neutralising antibodies is not clear, an increased frequency of such antibodies for the biosimilar product relative to the reference medicinal product would contradict the concept of biosimilarity. However, lower immunogenicity alone would have to be explained but may not preclude biosimilarity if efficacy is shown to be comparable in the various categories of patients according to their immune response (as previously defined) and provided all other data (quality, non-clinical, PK, PD, and safety) are supportive of biosimilarity.*
Lines 208-210	3	Comment: The concept of biosimilarity would also be contradicted, if neutralizing antibodies would have a substantial effect on the pharmacodynamic effect of IFNB or on the outcome measures of the clinical study (see general comment 2B&C).	Partly accepted. *It is known that development of NAbs has an effect on treatment efficacy, thus an impact of NAbs on PD/efficacy is indeed expected. The concept of biosimilarity would only be contradicted if this impact differs between the biosimilar and reference product.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change: Although the clinical impact of binding, non-neutralising antibodies is not clear, an increased frequency of such antibodies for the test product relative to the reference product would contradict the concept of biosimilarity. Moreover, it would be of concern, if neutralizing antibodies have an substantial impact on the pharmacodynamic effect of the biosimilar or on the outcomes of the clinical study.	*The sentence on the binding non-neutralising antibodies only means that, although their impact is less clear, it is still considered important that their incidence is not higher with the biosimilar compared to the reference.* *A clarification statement has been added.* *<u>New wording</u>* *The immune response to the biosimilar and reference medicinal products is expected to be comparable with regard to the incidence and titres of antibodies (neutralising or not) as well as their impact on efficacy; although the clinical impact of binding, non-neutralising antibodies is not clear, an increased frequency of such antibodies for the biosimilar product relative to the reference medicinal product would contradict the concept of biosimilarity. However, lower immunogenicity alone would have to be explained but may not preclude biosimilarity if efficacy is shown to be comparable in the various categories of patients according to their immune response (as previously defined) and provided all other data (quality, non-clinical, PK, PD, and safety) are supportive of biosimilarity.*
Lines 210-211	2	Comment: It seems unclear why there seems to be clarity in the concept that NAbs can impact clinical outcome, but then its evaluation is postponed as part of a post-marketing commitment-type effort. If a "biosimilar" has a high incidence of NAbs (i.e., >50%). It would be difficult for an agency to approve such a product and be sufficiently reassured that the post-	Partly accepted. *It is obvious that a higher incidence of NAbs with the biosimilar compared to the reference product would preclude a conclusion that they are comparable with regard to their immunogenicity. No post-authorisation commitment would be able to change this difference; the products would not be*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		marketing commitment will address any concerns on clinical outcomes and efficacy. Proposed change (if any): Add after line 211. "The final measurements could be obtained as part of a post-marketing commitment where there is *reasonable* expectation of similar clinical outcomes based on NAb titre levels during the clinical trial."	*considered biosimilar.* *The additional 6-month data, which can be provided post-authorisation, are only required to better characterise the impact of Nabs on efficacy outcomes.* *This has been further clarified.* *New wording:* *MRI activity and clinical relapses should be compared between these categories for both the biosimilar and reference product. The impact of NAbs on clinical outcomes is unlikely to be sufficiently ascertainable before 12 months of therapy and thus will need to be further evaluated post-authorisation as part of the risk management plan.* *The immune response to the biosimilar and reference medicinal products is expected to be comparable with regard to the incidence and titres of antibodies (neutralising or not) as well as their impact on efficacy; although the clinical impact of binding, non-neutralising antibodies is not clear, an increased frequency of such antibodies for the biosimilar product relative to the reference medicinal product would contradict the concept of biosimilarity.*
Lines 213-217	2	Comment: "The risk management plan should particularly focus on rare events such as autoimmune disorders and on the potential effects of unwanted immunogenicity". The guideline could be more specific on the type of	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		adverse events expected to be linked to immunogenicity. Proposed change (if any): …It should be based on the known identified and potential risks of the reference product as described in its risk management plan or product information. The risk management plan should further ~~particular~~ focus on rare events such as autoimmune disorders, important missing information such as pregnancy outcome and on the potential effects of unwanted immunogenicity.	Accepted. *New wording:* *The risk management plan should further focus on rare events such as autoimmune disorders, on adverse events of special interest such as hepatotoxicity and depression, on the potential effects of unwanted immunogenicity, and on important missing information such as pregnancy outcome (a pregnancy registry for IFN-β containing products is mandatory).*
Line 213-218	4	Comment: The working party could consider including a statement on the existence of the pregnancy registry for interferon beta containing products, which was requested by the CHMP in April 2006.	Accepted. *New wording:* *The risk management plan should further focus on rare events such as autoimmune disorders, on adverse events of special interest such as hepatotoxicity and depression, on the potential effects of unwanted immunogenicity, and on important missing information such as pregnancy outcome (a pregnancy registry for IFN-β containing products is mandatory).*
Lines 214 to 217	3	Proposed addition: …It should be based on the known identified and potential risks of the reference product as described in its risk management plan or product information. The risk management plan should further ~~particular~~ focus on rare events such as autoimmune	Accepted. *New wording:* *The risk management plan should further focus on rare events such as autoimmune disorders, on adverse events of*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		disorders, important missing information such as pregnancy outcome and on the potential effects of unwanted immunogenicity.	*special interest such as hepatotoxicity and depression, on the potential effects of unwanted immunogenicity, and on important missing information such as pregnancy outcome (a pregnancy registry for IFN-β containing products is mandatory).*
Line 217-218	4	Comment: Apart from the obligatory pregnancy registry, the RMPs of the currently authorised interferon beta containing (reference) products only include routine pharmacovigilance activities. The draft guideline however only mentions additional pharmacovigilance activities for monitoring *rare events (such as autoimmune disorders),* and *potential effects of unwanted immunogenicity*. Though additional pharmacovigilance activities are encouraged to monitor unwanted effects of immunogenicity (given the high incidence of neutralising antibodies, and the possibility of an altered immunogenic profile of the biosimilar as compared to the reference product), routine pharmacovigilance may in many cases be sufficient for rare events. Inclusion of a statement on the possibility of routine pharmacovigilance is therefore encouraged.	Accepted. *New wording* *This could be managed through routine pharmacovigilance, the extension of the pre-authorisation trial (in particular regarding immunogenicity as previously mentioned), a dedicated observational study or the participation in an existing registry.*
Line 218	2	Comment: In order to ensure traceability at the level of adverse events reporting and general pharmacovigilance of the similar biological medicinal products in respect to the reference products it is suggested that a suffix is added to the INN for new	Not accepted. *This is outside the remit of the CHMP/EMA.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		products. This will allow proper pharmacovigilance. Proposed change (if any): Add after line 218 "Similar biological medicinal products will include a suffix in their INN in order to ensure adequate pharmacovigilance."	
Lines 219 - 222	2	Comment: It is well understood that relapsing MS becomes more difficult to treat in the more advanced stages. A 5% higher efficacy on relapses could make the difference between a drug that is effective in EDSS>3.5 and one that is not, even though they may share the same MOA. Additionally, the mechanism of action of IFNs is not well understood and different mechanisms could be involved, given the different patterns of the disease, at the different stages of MS. As outlined in CPMP/EWP/561/98 Rev 1, efficacy in patients with relapsing-remitting MS cannot be extrapolated to patients with purely progressive forms of MS, i.e. patients with primary or secondary progressive MS without superimposed relapses. Therefore, demonstration of efficacy and safety in confirmed RRMS should not allow extrapolation, without supportive data, to non-relapsing forms of MS (e.g. SPMS) or to the earlier stage of clinically isolated syndrome (CIS) prior to definite MS diagnosis.	Partly accepted. *No IFN-β has an indication in non-relapsing forms of MS and it is agreed that IFN-β is only active on relapses. One of the IFN-β products has an indication in SPMS but only for patients suffering relapses.* *Patients with a single clinical demyelinating event at risk of developing MS (the risk being determined based on the MRI picture) would nowadays be diagnosed as definite RRMS according to the revised McDonald criteria (2010).* *Extrapolation is not only based on the clinical trial but on the entire data package including quality, non-clinical, PK & PD, as well as efficacy/safety in RRMS. If all these aspects support similarity, it is considered that the other RMS indications can be granted without additional study.* *It has been clarified that only RMS is concerned.* *New wording:* *Although not precisely understood, the mechanism of action*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change: ~~Although not precisely understood, the mechanism of action of IFN-β can reasonably be assumed to be the same whatever the stage of MS. Therefore, the demonstration of efficacy and safety in confirmed RRMS will allow extrapolation to other indications of the reference medicinal product in MS.~~ Ideally the patient population treated with the drug should be well matched with the study population in which the drug was tested. Efficacy in patients with relapsing-remitting MS cannot be extrapolated to patients with purely progressive forms of MS, i.e. patients with primary progressive MS or secondary progressive MS without superimposed relapses. Treatments focused on modifying relapses may be useful in patients with secondary progressive MS still suffering relapses. This situation is covered under the indication of treatment for relapsing MS. However, it must be taken into account that some effect on relapses without an accompanying effect on disability may be considered, in these patients, less important than in RRMS. Thus a certain efficacy seen in RRMS cannot be entirely extrapolated to patients with SPMS even if they still have relapses. On the other hand as long as safety and immunogenicity are not entirely understood (and they cannot be - especially as far as rare adverse events are concerned - after only a short study period) one might want to carefully consider the	*of IFN-β on MS relapses can reasonably be assumed to be the same whatever the stage of RMS. Therefore, demonstration of efficacy and safety in confirmed RRMS will allow extrapolation to the other indications of the reference medicinal product in RMS.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		risk/benefit balance before extrapolating from RRMS studies only and administering it to CIS patients without specific study data in this population or long-term safety data.	
Lines 220-222	3	Comment: As outlined in CPMP/EWP/561/98 Rev 1, efficacy in patients with relapsing-remitting MS cannot be extrapolated to patients with purely progressive forms of MS, i.e. patients with primary or secondary progressive MS without superimposed relapses. Proposed change: Ideally the patient population treated with the drug should be well matched with the study population in which the drug was tested. Efficacy in patients with relapsing-remitting MS cannot be extrapolated to patients with purely progressive forms of MS, i.e. patients with primary progressive MS or secondary progressive MS without superimposed relapses. Treatments focused on modifying relapses may be useful in patients with secondary progressive MS still suffering relapses. This situation is covered under the indication of treatment for relapsing MS. However, it must be taken into account that some effect on relapses without an accompanying effect on disability may be considered, in these patients, less important than in RRMS. Thus a certain efficacy seen in RRMS cannot be entirely extrapolated to patients with SPMS	Not accepted. *Whatever the plethora of downstream events triggered by interferon, the common starting point is the binding of interferon to its receptor, which is considered the critical step for extrapolation purposes.* *Furthermore, extrapolation is based on the entire data package including quality, non-clinical, PK & PD, as well as efficacy/safety in RRMS. If all these aspects support similarity, it is considered that the other MS indications can be granted without additional study.* *Of note, no IFN-β product has an indication for non-relapsing forms of MS.* *<u>New wording</u>:* *Extrapolation of clinical efficacy and safety in confirmed RRMS to the other indications of the reference medicinal product in MS is possible on the basis of the totality of the evidence provided from the comparability exercise.*

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		even if they still have relapses. On the other hand as long as safety and immunogenicity are not entirely understood (and they cannot be - especially as far as rare adverse events are concerned - after only a short study period) one might want to carefully consider the risk/benefit balance before extrapolating from RRMS studies only and administering it to CIS patients without specific study data in this population or long-term safety data.	