Description of problem	Description of proposed amendment	Justification for amendment	ERG response
In several instances in the ERG report, it is suggested that there was no consistent strategy about the selection of NMA scenarios for the CS economic model. In page 76, second paragraph, the ERG indicates that there might be intentional selection of scenarios to favour nintedanib: *This creates an impression (potentially falsely) that scenario analyses may have been tried until one was found that provided a favourable result.*	In each instance, the report should include the justification provided in the CS, for the reader to judge whether this is sufficient or not.	The ERG correctly points out that the all-evidence scenario was not always used in the model. However, it is not clarified why they believe that this should be the case. Based on Boehringer Ingelheim's assessment of study (or outcome) biases a sensitivity analysis was performed by excluding studies. This creates a list of options for the model. The CS provides sufficient justification for the chosen scenario; for instance the CS in p174 describes the reasons S3 was selected for acute exacerbations; heterogeneity and Japanese population. Furthermore, the selection was not based on the result for the NDB cost-effectiveness. If that was the case, then for OS (the most important driver for the cost-effectiveness results) scenario 2 or 3 would have been selected as these are more favourable for nintedanib than the base case scenario. For this reason, the indication that Boehringer Ingelheim may have intentionally selected scenarios in favour of nintednib is factually inaccurate and misleading. Additionally, the message that there is no consistent strategy employed when choosing NMA scenarios is factually inaccurate and should at least be accompanied by the reason included in the CS to allow the reader to make an informed decision.	Not a factual inaccuracy As stated on p.76 (now p.75) of the ERG report, no overarching logic for the scenario analyses conducted by the company was described. Decisions about which scenario analyses to undertake appeared to have been made post-hoc. The ERGs concern in this section of the ERG report is primarily about which scenario analyses were undertaken for each outcome rather than about the process of choosing which scenario to use from those that are reported. As Table 7 of the ERG report shows (ERG report p. 46-47, now p. 45-46) the scenario analyses undertaken differed across the different outcomes. For example, the ASCEND study (King et al.) contributed evidence to four NMA outcomes: loss of lung function, overall survival, overall discontinuation, discontinuation due to AEs. For two outcomes a scenario analysis was conducted that excluded only the ASCEND study (loss of lung function scenario 2, overall discontinuation scenario 4). The CS does not indicate why the same scenario was not reported for the other two outcomes (i.e. excluding only ASCEND from the overall survival and discontinuation due to AEs outcomes).
Page 57: *However, the CS does not explain whether*	Delete this sentence	This statement is incorrect. In page 105 of the	Noted. The sentence has been removed. A similar sentence in paragraph 3 of page 57

this includes data from only the licensed dose and placebo arms of the TOMORROW trial or from the full study (which included study arms with unlicensed doses).		CS it is stated that: *With reference to nintedanib and pirfenidone treatments, the analysis considered the licensed dose of each treatment: 150mg BID for nintedanib and 2,403mg daily for pirfenidone. The alternative treatment arms that investigated lower doses of nintedanib (50mg and 100mg) or pirfenidone (1,197mg and 1,200mg) were not used in the analysis. Sensitivity analysis with inclusion of the lower dose arms was not considered necessary given the available data from the high dose arms and the potential impact of the treatment dose to the relative effect.*	(now page 56) has also been removed. Also a similar sentence in paragraph 2 on page 58 (now page 57) has been removed.
Page 62: *The NMA output for the nintedanib vs. pirfenidone comparison in the all evidence scenario (fixed effect) was a median OR of 0.96 (95% CrI 0.36 to 2.85; CS Table 55 p. 120)*	Change the UL of the CrI to 2.58	The upper limit of the CrI is incorrect.	This is a typographical error. The UL has been corrected to 2.58 on page 62 (now page 61)
In several instances the ERG report mentions the differences in study duration.	Complete all mentions of the possible bias in the NMA due to the study duration, with the information obtained by the ERG's clinical expert.	Although clinical advice to the ERG suggests that a difference of 20 weeks is not long enough to observe any differences, this is not always repeated.	This is not a factual inaccuracy, no change required. Clinical advice was that a 20 week difference in study duration might be too short to observe a difference in mortality or a difference in FVC outcomes and this advice is presented in the ERG report where appropriate. No other information was obtained from the clinical expert regarding other outcomes.
Page 117: Analysis 4 description	Correct to initial decrement to 0.14. Boehringer Ingleheim are uncertain about the ERG's	Possible typographical error. Please clarify the calculation in the ERG analysis. How did the ERG derive the difference for the new	This is a typographical error which has been corrected to 0.14 where it appears on p. 15 (now p.14) and pp. 116-118 (now pp.

	calculations for this analysis	exacerbation cases?	115-117).
Page 98: *Secondly, although the company presented results from both the fixed effect and random effects models in the NMA, the company chose estimates from the fixed effect models across all the outcomes to use in the economic model despite the clinical evidence suggesting that random effects models performed better for acute exacerbations and serious cardiac events for the all evidence scenario of the NMA.*	Delete this sentence	It is true that the all-evidence scenario was not always used in the model and that the random effects model performed better than the fixed effects model in the all-evidence scenarios of acute exacerbations and serious cardiac events. However, in those cases high levels of heterogeneity in the results were observed. For acute exacerbations, in pages 171-172 of the CS it is stated that: *The NMA scenario used in the model excluded Azuma et al. 2005, Taniguchi et al. 2010, and Homma et al. 2012 (See Appendix B, Scenario 3) because inclusion of these studies introduced a high level of heterogeneity to the other PFN and NAC studies. Furthermore, the fact that the above mentioned studies were Japanese-only trials with different treatment dosages than the European and US studies supported their exclusion from the base-case analysis. The fixed effects model was used as it performed better than the random effects model.* For serious cardiac events, in page 179 of the CS it is stated that: *The NMA scenario implemented in the model for serious cardiac events excluded Richeldi et al. (See Appendix B, Scenario 2) because inclusion of this study introduced a high level of heterogeneity when compared with Richeldi et al. 2014.* In the scenarios selected for acute	The ERG did not state that random effects was preferable to fixed effect in the acute exacerbations (Scenario 3) and serious cardiac events (Scenario 2) evidence scenarios. It was pointed out that in the all evidence scenario the random effects model had performed better than the fixed effect model that was used in the model. This was the rationale for the ERG to conduct a scenario analysis using the all evidence scenario along with using both fixed and random effects, for all outcomes. For clarity the text on p.98 (now p.97) of the ERG report has altered and now reads "Secondly, although the company presented results from both the fixed effect and random effects models in the NMA, the company chose estimates from the fixed effect models across all the outcomes to use in the economic model (as favoured in the individual evidence scenarios used) despite the clinical evidence suggesting that random effects models performed better for acute exacerbations and serious cardiac events for the all evidence scenario of the NMA."

		exacerbations (S3) and serious cardiac events (S2) the fixed effects model performed better than the random effects one. It is factually inaccurate to state that the random effects model is preferable in these scenarios.	
Pg 17 *The SmPC does not state the length of treatment, but the company suggests in CS Table 5 (CS p. 25) that nintedanib should be used until disease progression or the need to discontinue due to unacceptable adverse events.*	Remove sentence or replace with: The SmPC does not state the length of treatment.	Boehringer Ingelheim in no-way intended to suggest that treatment should be discontinued following 'disease progression'. Efficacy data demonstrates that nintedanib slows disease progression, not that it halts or delays it. It is expected that a patient's disease will progress on treatment, but to do so more slowly than those not treated. There is no data to support stopping treatment in response to any specified marker of 'disease progression' and this advice is not given in the SmPC.	Not a factual inaccuracy, no change needed. CS Table 5 (CS p.25) clearly states in two rows (Anticipated average interval between courses of treatments and Anticipated number of repeat courses of treatment) "Continuous daily treatment until disease progression or unacceptable adverse events".
Pg 21 *Clinical expert advice to the ERG is that, approximately, a FVC < 80% predicted indicates mild IPF.*	Change to: Clinical expert advice to the ERG is that, approximately, a FVC > 80% predicted indicates mild IPF.	Presumed typographical error using "less than" sign rather than "more than".	This was indeed a typographical error on ERG report p.21 (now p.20) which has now been corrected.
Pg 14, Pg 121 *The population used in the economic model may not represent the clinical population* *treated in the UK because they have included patients with FVC% predicted more than* *80% which represents IPF that is milder than would typically be seen in current UK* *practice.*	Change to: The population used in the economic model may not represent the clinical population treated in the UK because they have included patients with FVC% predicted more than 80% which represents IPF that is milder than reported in the description of UK practice in the pirfenidone STA, though this is not a universally held view.	Boehringer Ingelheim do not believe that this description of the current clinical population is accurate, though admit that it as described in the pirfenidone STA. Boehringer Ingelheim have clinical database data, data from the UK named patient programme use of nintedanib and external expert advice that contradicts this opinion. In an area where there was no NICE approved treatment prior to 2013, it may be inappropriate to base an economic model on historical management approaches. It is a logical necessity in a progressive disease that most patients must have begun the IPF pathological process with an FVC above 80%, and thus it is factually inaccurate to suggest the majority of	The ERG does not believe this is a factual inaccuracy but concedes that the wording on ERG report p.14 (now p.13) could have been improved. The ERG has amended this text to read "The population used in the economic model may not fully represent the clinical population treated in the UK".

		UK patients have an FVC% predicted of <80%. Given the sub group analyses from INPULSIS that show a rate of similar rate of decline in the placebo arms and similar treatment effect regardless of baseline FVC, it seems illogical to restrict an economical model to those with an FVC of <80%.	
Pg 21 - *In their clarifications response, the company indicated that subgroup analyses using a FVC% predicted threshold of 80% have also been conducted and published. The company referred to an analysis published in "Maher et al. ERS 2015" but did not provide a full reference for this source. The ERG was unable to locate this reference and therefore was not able to check the analyses and results provided in it. The ERG notes that results for the 80% threshold subgroup analyses are not presented in the CS.*		This is an error on Boehringer Ingelheim's behalf. The abstract is due for presentation at the European Respiratory Society annual conference in September. It can be supplied now as academic in confidence data on file.	Thank you, noted