

Round 2 reasons for change

Participants were encouraged to report the reason why their rating changed from Round 1. This was only necessary when the rating crossed a boundary (boundaries as follows: (1-3), (4-6), (7-9)).

Outcome	Round 1 rating	Round 2 rating	Reason
[37] SPIRIT: Likely impact of potential interactions on interpretation	6	7	Based on other participants ratings
	6	7	Was borderline previously - I think this is important to consider at SPIRIT stage
	3	6	Considered the response of others; researched the topic more thoroughly since the first occasion; and placed even greater emphasis on consistency of responses.
	8	6	on re-reading i judged slightly less important
	7	6	It does not appear critical; therefore adjusted (including in light of others' responses)
	7	6	I agreed with the comment that this is important for the publication but less so for potential impact in the protocol
	9	6	Actually I think this is covered by a later item which asks for basically the same thing
	7	5	Important in interpretation but not so for the protocol
[38] CONSORT: Likely impact of identified interactions on interpretation	3	9	Considered the response of others; researched the topic more thoroughly since the first occasion; and placed even greater emphasis on consistency of responses.
	6	7	Although rating was upgraded; don't think we need 38 and 48 as separate items
	6	8	was unsure last time but seen other voting
	6	7	Based on other participants ratings
	5	9	Based on further reflection
	4	8	Helps with interpretation/serves as sensitivity thought experiment
	6	7	useful for reader
	7	6	on re-reading i judged slightly less important
	7	6	It does not appear critical; therefore adjusted (including in light of others' responses)
9	6	Actually I think this is covered by a later item which asks for basically the same thing	
[49] CONSORT: Whether adherence to intervention might have been affected by inclusion of other factors	1	4	Again I can see why this might be important in some trials. I am not sure that it should be mandatory.
	6	7	Based on other participants ratings
	7	6	reconsidered after seeing others's responses
	9	6	I agreed with the comment that I'm not sure how you might reliably tell

[50] CONSORT: Effect of multiplicity of analyses (if relevant)	6	7	Based on other participants ratings
	6	7	I remain on the fence; but can see that the issue is important and better discussion of multiplicity would improve trial reporting.
	6	7	Not sure why I put 6 initially. I am interpreting this now in the context of type I error control. Hence the 7
	2	4	Too dismissive of this before.
	4	7	towards modal response of other respondents
	10	6	last time felt i did not understand the item but on reflection realise I do
	7	6	On reflection; the effect of such multiplicity is usually obvious so perhaps not as important
	9	6	In hindsight I felt this was important but not critical as it may not always be relevant
	7	6	again more about theorizing