PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Do consultants do what they say they do? Observational study of the extent to which clinicians involve their patients in the decision-making process.
AUTHORS	Driever, Ellen; Stiggelbout, Anne; Brand, Paul

VERSION 1 – REVIEW

REVIEWER	Carmona, Chris The University of Sheffield
REVIEW RETURNED	24-Aug-2021

GENERAL COMMENTS	Thanks for undertaking this study. It is interesting to see how little agreement there is between consultants own perceptions of their decision making style and the reality of it using a validated instrument. Especially interesting to note that the 'worst culprits' are individuals who report their decision making style as SDM. It is also sobering to note how poorly consultants score on SDM with no-one scoring over 40/100. Overall, I think the paper makes a useful contribution and highlights a problem for the implementation of SDM that we know about but haven't resolved, ie, that most people think they are doing SDM but actually aren't, and on that basis I think it is worth publishing. I would suggest a few amendments detailed below: Key points - paper needs more clarity about OPTION5 scores. It begins saying each parameter is scored 0-4, giving a total score 0-20, which is scaled up on to a % scale. Then, under the 'OPTION5 scores' header (p10/23) it presents the scores scaled up, so out of 20 per parameter/item rather than out of 20 for all 5 parameters/items. In the next paragraph, it presents the mean OPTION5 score for the main decision as 16.8. I assume this is 16.8 out of 100 rather than 16.8 out of 20. It would really help to be very clear about the scale used each time you present these scores. Minor points - ensure consistency of reference to OPTION5 with the 5 either consistently superscript or consistently not. - p7/23 line 3 reliably rather than reliable? - 8/23 line 15-17 sentence grammatically poor. Suggest delete 'taken'? - 10/23 line 15 "After excluding 36 consultations FROM THE analysis..." - 11/23 line 59 Perhaps rephrase 'unconsciously incompetent' (also in conclusion) - it sounds a little harsh?

REVIEWER	Johnson, Rachel University of Bristol, Population Health Sciences
REVIEW RETURNED	13-Sep-2021

GENERAL COMMENTS	This study addresses a useful question (the extent to which clinicians' self-report of shared decision making agrees with observer-measured decision making), its mostly very clearly described, well written, and has some useful results. However, there are some methods details that are not currently clear, and these need to be addressed before publication. Firstly the title - I don't think this is an analysis of patient involvement, because this study doesn't report any patient perspectives, and the option 5 asks about clinician behaviour. Would shared decision making be better? Abstract: states the objective as looking at treatment decisions but the analysis is not restricted to treatment decisions Methods – I read the main paper before the abstract and the control preferences scale came as a surprise – isn't that what was used in a previous study? And the CPS is not mentioned at all in the main text unless I missed it? Suggest this needs to be more clear, and it should be described in the main methods section. in the methods it says 'Shared decisions making' so needs the extra s edited out. Intro: Well written. a brief intro to the different ways to assess SDM would be helpful, there are many measures, the choice of Option 5 should be justified. Methods The previous paper using the CPS is referenced, but the methods need more information about that study in order to understand this one without reading the references. Including the use of the CPS. Here it says that different medical consultants were included, later it says medical, surgical and 'supportive' (I'm not sure what that term means) but in the results only medical and surgical again - can the terms be made consistent please? Why consultants only? what about more junior doctors? Relationship to this study to the previous one is not very clear at the moment – how were the participants sampled for that study (presently there is no detail) , and how were they sampled for this study? who declined? Is the potential sample representative ? What specialities were covered? Was the CPS used? Because figure 1 reports a self-reported Option 5 scale? When was the cross-sectional study in relation to this? if clinicians self-reported using option 5, did they have access to recorded consults to do this? (as its an observed scale?) if they did, this has different implications - and would likely have affected how they consulted? if they scored it without recorded consults, is this valid? When was the self-report in relation to the observed consults? 10 encounters with different patients? How did you determine this? Why did you only focus on the clinician? Suggests SDM is a clinician behaviour - isn't it a shared process? the Option 5 scale also reports clinician behaviour - worth saying why this was chosen. Please say who the people were who were coding the decision making (what is their background, relationship to participants).

	How was the 'main decision arising from the chief complaint' determined? Eg from whose perspective? Often there are multiple 'complaints' particularly in some specialities – what then? Who decided? Later in results evident that more than one decision was chosen for lots of consults - so how did you decide which to include and who made this decision? More detail about this please. What did the participants know about what the study was about? Did they know how you would be judging them - what scale would be used? minor point: Missing bracket page 9 line 24. Results You have stated that the evidence that non-parametric and parametric analyses were similar, so only parametric presented - I think we need evidence of this otherwise we are taking it on trust. What if the same decision is covered over a series of consults? And some components of SDM might have happened in previous consultations? Needs consistency in how decisions are referred to -at times refers to 'main decision' if there were <1 in a consult, which was the main one? P 11 lines 14-22 – comments from a non-statistician here! - describing comparison between consultants who reported self as paternalistic etc – what is the method for this 3 way comparison (don't think its covered in methods?) , what does the p value mean here? is it meaningful? Is the choice of the mean for comparison meaningful if there is big variation between consults? Analysis of relationship between reported decision style and OPTION score – is this the mean score again? not clear Higher option scores 'only significantly related to' category of decisions' etc – not really clear what direction the associations described are in – which had higher scores? I think this needs to be stated in the main results as well as in the display items. Discussion Starts by saying that this paper looks at 'routine decisions' it would be helpful to understand more about the type of decisions being looked at. I think there are other possible explanations for the results / strengths and weaknesses to consider. This could include, for example, that there were decisions that were distributed over a series of consultations, and that only parts of the SDM process were seen in each consultation. or that from the clinician perspective, SDM means something different to that captured in the option scale (we know that there is not a great deal of agreement across different SDM scales) Would more recently qualified consultants or junior doctors be better? Does this reflect a more recent focus on SDM training? Would a SDM measure from eg. a patient perpective have different results? You describe the decisions as 'routine' - is SDM more likely for less 'routine' decisions? Did you categorise the key decisions correctly?

VERSION 1 – AUTHOR RESPONSE

Reviewer 1 (dr. Chris Carmona, The university of Sheffield)
Thanks for undertaking this study. It is interesting to see how little agreement there is between consultants own perceptions of their decision making style and the reality of it using a validated instrument. Especially interesting to note that the 'worst culprits' are individuals who report their decision making style as SDM. It is also sobering to note how poorly consultants score on SDM with no-one scoring over 40/100.

Overall, I think the paper makes a useful contribution and highlights a problem for the implementation of SDM that we know about but haven't resolved, ie, that most people think they are doing SDM but actually aren't, and on that basis I think it is worth publishing. I would suggest a few amendments detailed below:

C	Comment	How is this comment addressed?	Where can the change be found?
1	The paper needs more clarity about OPTION5 scores. It begins saying each parameter is scored 0-4, giving a total score 0-20, which is scaled up on to a % scale. Then, under the 'OPTION5 scores' header (p10/23) it presents the scores scaled up, so out of 20 per parameter/item rather than out of 20 for all 5 parameters/items. In the next paragraph, it presents the mean OPTION5 score for the main decision as 16.8. I assume this is 16.8 out of 100 rather than 16.8 out of 20. It would really help to be very clear about the scale used each time you present these scores.	We clarified the description of the OPTION5 score and the presentation of its results, as suggested.	Results – page 9 and 10
2	Ensure consistency of reference to OPTION5 with the 5 either consistently superscript or consistently not.	We changed this in the text, with the 5 consistently in superscript.	Throughout the text
3	7/23 line 3 reliably rather than reliable?	We changed this.	
4	8/23 line 15-17 sentence grammatically poor. Suggest delete 'taken'?	We changed this.	Methods - Page 8
5	10/23 line 15 "After excluding 36 consultations FROM THE analysis..."	We changed this.	
6	11/23 line 59 Perhaps rephrase 'unconsciously incompetent' (also in conclusion) - it sounds a little harsh?	We understand that this may come across a little harsh, but we do believe that it is right term to use. The term was derived from a theoretical framework in medical	Discussion

		education, which we described in more detail in the revised discussion section.	

Reviewer 2: (dr Rachel Johnson, University of Bristol)

This study addresses a useful question (the extent to which clinicians' self-report of shared decision making agrees with observer-measured decision making), its mostly very clearly described, well written, and has some useful results. However, there are some methods details that are not currently clear, and these need to be addressed before publication.

C	Comment	How is this comment addressed?	Where can the change be found?
1	Firstly the title - I don't think this is an analysis of patient involvement, because this study doesn't report any patient perspectives, and the option 5 asks about clinician behaviour. Would shared decision making be better?	According to the developer of OPTION5, the instrument assesses the degree of patient involvement by the provider (or by patient's own assertiveness). Although it is often used as a proxy measure of the degree of SDM, we believe we are most accurate when we describe it in the title as follows: *Do consultants do what they say they do? Observational study of the extent to which clinicians involve their patients in the decision-making process.*	Title page, and a lso in the abstract and introduction .
Abstract			
2	States the objective as looking at treatment decisions but the analysis is not restricted to treatment decisions	It is based on the main decisions, but those are not always treatment decisions. We have therefore deleted the word 'treatment'.	Abstract – page 3
3	Methods – I read the main paper before the abstract and the control preferences scale came as a surprise – isn't that what was used in a	The sentence on page 7 referring to the previous study refers to the data collected by the modified CPS in the previous study, but we did not mention this as such in manuscript. In the revised version, we changed the text to specifically describe the use of the modified CPS.	Methods – page 7

	previous study? And the CPS is not mentioned at all in the main text unless I missed it? Suggest this needs to be more clear, and it should be described in the main methods section.		
4	In the methods it says 'Shared decisions making' so needs the extra s edited out.	We deleted the extra s.	Abstract - Page 3
Introduction			
5	Well written. a brief intro to the different ways to assess SDM would be helpful, there are many measures, the choice of Option 5 should be justified.	Indeed, there are different ways to assess SDM. When we designed this research, we deliberately chose the OPTION5, for several reasons. We added this to the methods section instead of the introduction.	Methode – page 7
Methods			
6	The previous paper using the CPS is referenced, but the methods need more information about that study in order to understand this one without reading the references. Inclu ding the use of the CPS.	We changed the text to make it more clear.	Methods – page 6
7	Here it says that different medical consultants were	This was a mistake, for which we apologise. Originally, we intended to classify participants into medical, surgical and supportive disciplines. Because there were only 2	Methods – page 8.

	included, later it says medical, surgical and 'supportive' (I'm not sure what that term means) but in the results only medical and surgical again - can the terms be made consistent please? Why consultants only? what about more junior doctors?	consultants from a supportive discipline in the study sample (both anesthesiologists), we refrained from analyzing these two as a separate group, and classified them as a medical discipline instead. We changed the text to describe this accurately and consistently. In this study, we only included consultants. We agree that it would also be useful to analyse video recordings of consultations of junior doctors. This could be added as a suggestion for future research. We think we are now clear and consistent in our description of study participants.	
8	Relationship to this study to the previous one is not very clear at the moment – how were the participants sampled for that study (presently there is no detail) , and how were they sampled for this study? who declined? Is the potential sample representative ? What specialities were covered? Was the CPS used?	We extended the description of the recruitment procedure and we added information on the consultants' disciplines in Table A in a supplementary appendix. We believe our sample is representative of medical specialist practice in a general teaching hospital, and recognize that further studies are needed in other settings and countries to increase the generalizability of our findings.	
9	Because figure 1 reports a self-reported Option 5 scale? When was the cross-sectional study in relation to this? if clinicians self-reported using option 5, did they have access to recorded consults to do this? (as its an observed scale?) if they did, this	Figure 1 shows the mean OPTION5 score of 41 consultants categorized by the self-reported usual decision-making style (assessed with the modified CPS). We added extra information about the CPS in the Methods section and we changed the title of figure 1.	Methods – page 7 Figure 1 – title

	has different implications - and would likely have affected how they consulted? if they scored it without recorded consults, is this valid? When was the self-report in relation to the observed consults?		
10	10 encounters with different patients? How did you determine this?	The researchers who conducted the data collection (EMD and RH) made sure that every patient who gave informed consent to video-tape the encounter with the consultants, was unique.	No changes made
11	Why did you only focus on the clinician? Suggests SDM is a clinician behaviour - isn't it a shared process? the Option 5 scale also reports clinician behaviour - worth saying why this was chosen.	The OPTION5 instrument is the most commonly used instrument to assess the degree of patient involvement in decision making. The instrument focuses on clinician behaviour, which fitted our research question in which we wanted to assess whether medical consultants do what they say they do in terms of decision making in medical consultations. We added information on why we chose the OPTION5 instrument to the Methods section.	
12	Please say who the people were who were coding the decision making (what is their background, relationship to participants).	We described in the Methods section (under recruitment of participants) what the relationship was between the participants and members of the research team. In the paragraph about the OPTION5 instrument, we described the background of EMD and RH, the researchers who scored the encounters.	Methods – page 6 and 7
13	How was the 'main decision arising from the chief complaint' determined? Eg from whose perspective? Often there are multiple	We added information about this process, hopefully it is now more clear.	Methods – page 7

	'complaints' particularly in some specialities – what then? Who decided? Later in results evident that more than one decision was chosen for lots of consults - so how did you decide which to include and who made this decision? More detail about this please.		
14	What did the participants know about what the study was about? Did they know how you would be judging them - what scale would be used?	The participants were aware of the subject of the research: medical decision making. They knew we would look at their decision-making behaviour. Although this may have led to socially desirable behaviour, this would mean that the OPTION5 scores in real practice would be even lower than the ones we recorded in our study. The consultants were unaware which method we used to assess their decision-making behaviour. This issue is discussed as a limitation of the study in the Discussion section.	We added one sentence in strengths and limitation section – page 12
15	Minor point: Missing bracket page 9 line 24.	We added the bracket.	Page 9.
Results			
16	You have stated that the evidence that non-parametric and parametric analyses were similar, so only parametric presented - I think we need evidence of this otherwise we are taking it on trust.	For our main research question 'Do consultants do what they say they do?", we needed to compare the OPTION5 scores between consultants from three groups of self-reported decision-making style: paternalistic, shared or informative decision-making. Although these scores are not normally distributed by design, most studies in the literature have used parametric analyses (such as ANOVA) to analyse these data. For your information, the p value of the parametric ANOVA test comparing OPTION5 scores between the three groups was 0.017. With a non-parametric test (Kruskal Wallis), the p value of the difference between groups was 0.003, which we considered to be similar. We feel that this information is not essential for the interpretation of our study results, so we did not change it in the current revision of the manuscript.	No changes made.

		If the reviewer or editor feels that it is important that this information is added to the text, we would be happy to do so.	
17	What if the same decision is covered over a series of consults? And some components of SDM might have happened in previous consultations? Needs consistency in how decisions are referred to -at times refers to 'main decision' if there were <1 in a consult, which was the main one?	We clarified the description of the process of categorizing the decisions that we analysed. Thank you for this comment. We agree that some decisions may be spread over several consultations. This is not taken into account in the OPTION5 instrument, and may be one of the contributing factors explaining the low OPTION5 scores that we found. We added this as a limitation to the Discussion section.	Methods.
Discussion			
18	Starts by saying that this paper looks at 'routine decisions' it would be helpful to understand more about the type of decisions being looked at.	Because the word “routine” apparently caused confusion, we deleted it in the Highlights and in the strengths and limitations section.	Title page and discussion.
19	I think there are other possible explanations for the results / strengths and weaknesses to consider. This could include, for example, that there were decisions that were distributed over a series of consultations, and that only parts of the SDM process were	We agree with the reviewer and have now extended the description of the limitations.	Discussion – page 12

	seen in each consultation. or that from the clinician perspective, SDM means something different to that captured in the option scale (we know that there is not a great deal of agreement across different SDM scales)		
20	Would more recently qualified consultants or junior doctors be better? Does this reflect a more recent focus on SDM training?	We found no effect of the consultants' age on the OPTION5 scores. We deliberately limited ourselves to studying only consultants' behaviour in the present study. In a separate qualitative interview study among junior doctors, currently under review by a different journal, we explore registrars' perceptions of the decision-making process in medical consultations.	No changes made.
21	Would a SDM measure from eg. a patient perspective have different results?	We agree that this would be a worthwhile perspective, but feel it requires a completely different study to address this	No changes made.
22	You describe the decisions as 'routine' - is SDM more likely for less 'routine' decisions?	We deleted the word 'routine' (See comment 18), to avoid confusion. For example, we did not intend to suggest that SDM is more or less appropriate for routine decisions.	Deleted the word routine.
23	Did you categorise the key decisions correctly?	We added information about the categorization process to clarify this. We are confident that we categorized the key decisions correctly.	Methods – page

VERSION 2 – REVIEW

REVIEWER	Johnson, Rachel University of Bristol, Population Health Sciences
REVIEW RETURNED	04-Dec-2021

GENERAL COMMENTS	The paper is much improved and is very clear. It makes a clear contribution with well articulated implications and I think should be published. Only a few extremely minor edits: P57 line 8 there is a full stop missing.

	P 11 line 18 there is an extra full stop P12 line 19 Broadwell's and Maslow's – should be Broadwell and Maslow's P12 line 45 should read individual consultant's Line 49 – there are other possible factors other than patient, physician, organisational, e.g. the characteristics of the decision? P14 first paragraph misses full-stop at the end.