PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (http://bmjopen.bmj.com/site/about/resources/checklist.pdf) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Assessing medical student knowledge and attitudes about shared
	decision making across the curriculum: Protocol for an international
	online survey and stakeholder analysis
AUTHORS	Durand, Marie-Anne; Yen, Renata; Barr, Paul; Cochran, Nan; Aarts,
	Johanna; Legare, France; Reed, Malcolm; O'Malley, A. James;
	Scalia, Peter; Painchaud Guérard, Geneviève; Elwyn, Glyn

VERSION 1 - REVIEW

REVIEWER	Dr Aarti Bansal
	University of Shemeid. UK
REVIEW RETURNED	02-Feb-2017

GENERAL COMMENTS	Thank you for this well-written paper on a topical and important area. I have some queries. You state that there is only one study on medical students' attitudes towards SDM. I wonder if this demonstrates that you are looking at shared decision making too narrowly and in isolation. There have been studies looking at medical students patient-centred attitudes (one of which you refer to - Tsimtsiou, Zoi, 2007) using other scales, such as PPOS, which incorporate shared decision making. The reason I have answered no to study design is largely due a lack of clarity on your understanding of what constitutes shared decision making. Without this I was unable to make a judgment about whether the survey can answer the research question. I think it is important for the reader to understand the reasons for choosing the different clinical scenarios. Are you expecting students to answer differently based on whether the patient requires an urgent intervention or optional screening? Personally, as a practicing general practitioner, who is an advocate of shared decision making, I feel there would be a differential response. Perhaps something on your understanding of shared decision making (as this has been
	general practitioner, who is an advocate of shared decision making, I feel there would be a differential response. Perhaps something on your understanding of shared decision making (as this has been
	contested), would be useful. You write that "The information sheet intentionally does not mention SDM, but uses the term "health
	communication" to reduce potential respondent and desirability biases." However you do use the term "shared decision making" in
	your survey questions and I am not clear why. Thank you.

REVIEWER	Fania R. Gärtner
	Leiden University Medical Center
	Department of Medical Decision Making, the Netherlands
REVIEW RETURNED	06-Mar-2017

GENERAL COMMENTS	3. Is the study design appropriate to answer the research question?
	No, based on the collected data it is not clear to me how to realize

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

objective 3.: "Determine when and how to best deliver SDM training to medical students." On page 4 you state the objectives and also the aims of the Phase 2 data collection. It is not clear to me how these two exactly relate to each other. Also, it is not clear to me how the two phases relate to each other, are there different research aims for each phase or are they assumed to be analysed together to answer the same research questions?	
On page 6 you present the research questions. Based on the description of your survey I understand that you measure the following: demographic characteristics, attitude, knowledge, previous training, and based on the scenario's the availability of role models/ their view on the current clinical practice at their institution (If I understood it well) and their behavioural intention. I don't see how you will answer research question 3 and 4 with this data.	
4. Are the methods described sufficiently to allow the study to be repeated? No, what is the topic list?I do not fully understand your hypothesis about attitude and knowledge level across the curriculum.	
Analyses of qualitative data: Unfortunately I cannot judge about whether phase 2 will provide answers on the research questions, because I miss a topic list. Do you use the quantitative and qualitative data to answer the same research questions or different ones? I do not regard qualitative data suitable for all research questions. And if you use both phased to answer the same questions (I assume that you plan to do that because you use the word triangulation), how to you combine the data?	
Educational systems differ between countries and curricula differ within countries and also within one institution they differ between cohorts. I wonder if these differences allow to analyse all data together? I assume that interesting differences between the curricula/countries might be obscured. And how do you handle the different samples, students and experts in the analyses? Do you look for differences or do you analyse it as one set of data? In case of separate analyses, I wonder if 6-12 interviews is enough? You do check for data saturation, please state a pre-defined stopping rule?	
9. Do the results address the research question or objective? See point 3 and 4.	
11. Are the discussion and conclusions justified by the results. No, in the Discussion you state "We anticipate that the study findings will help determine the optimal content, format and timing of SDM training in medical education curricula worldwide." Based on the description of your study I am not convinced yet that you can realize this. I might miss some crucial information about your data collection that could convince me about that you indeed will realize these aims.	
12. Are the study limitations discussed adequately? No, a description of limitations is missing.	
Major comments	

Protocol What is the value of a paper describing the protocol for such study? I see the value for such papers in case of an RCT. Please convince me about the value of a publisched protocol for your study.
Relevance and validity The relevance of this study is not very clear to m. How exactly does your study help curriculum developers, teachers and students? Also, I expect that the evidence gained in this study is valid only shortly, because curricula change constantly.
Scenario's You describe clinical scenario's that you will use in your survey. I have doubts about their suitability. SDM is often described to be especially valuable for preference sensitive decisions, but in my opinion none of the three cases present a typical preference sensitive decisions. Two scenario's describe decisions about screening. Screening decision are often regarded to be suitable for informed decision making, which differs from SDM in some aspects.
Language Dutch students will receive the survey in English language and will be interviewed in English. Please explain why you do not make use of a Dutch survey? English language skills are overall good in the Netherlands but still vary between students. You have pilot tested the survey in Dutch students, but not in year 1 and 2. But these student's will be included in your sample, right? So why no pilot for these first two years for which Englisch language skills can be expected to be lower. I would prefer collecting data in the official language of the school/country. If not possible I would add the response option `I do not understand the question` to prevent any systematic error in the responses of the Dutch students, especially you have set up the survey with forced responses. For the survey you provide a glossary of terms for non-native English speakers, which might help to overcome language barrier. However, for the interviews such glossary does not work. Especially for qualitative data nuances and exact expression of respondents' opinions, experiences and feelings are very important, therefore I would recommend to interview the students in their own language.
The questionnaire I don't find Table 2 informative. I am not so much interested in the outline itself but much more in the content, which you present in the appendix. Please provide a reference for the attitude scale you use and the knowledge of SDM scale. Are these valid scales? And, please explain how attitude can be measured by the OPTION instrument, I know this scale only in the context of measuring the

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

Reviewer Name: Dr Aarti Bansal Institution and Country: University of Sheffield. UK Please state any competing interests: None declared

1. Thank you for this well-written paper on a topical and important area.

We thank the reviewer for the positive feedback.

2. I have some queries. You state that there is only one study on medical students' attitudes towards SDM. I wonder if this demonstrates that you are looking at shared decision making too narrowly and in isolation. There have been studies looking at medical students patient-centred attitudes (one of which you refer to - Tsimtsiou, Zoi, 2007) using other scales, such as PPOS, which incorporate shared decision making.

The reviewer raises an important point. Patient-centered care is typically understood and defined as 'the importance of better understanding the experience of illness and of addressing patients' needs' (Barry and Edgman-Levitan 2012, NEJM). The IOM defined patient-centered care as "care that is respectful of and responsive to individual patient preferences, needs, and values" and that ensures "that patient values guide all clinical decisions." Shared decision making focuses on one aspect of patient-centered care. According to Barry et al., it is defined as 'the process by which the optimal decision may be reached for a patient at a fateful health crossroads is called shared decision making and involves, at minimum, a clinician and the patient, although other members of the health care team or friends and family members may be invited to participate.' We have followed Barry et al.'s definition and therefore consider that although patient centered care is related to shared decision making, it may not always focus or include the process of shared decision making. We have made this clearer on page 4 of the revised manuscript:

Studies of the attitudes of doctors' in training towards patient-centered care suggest that patientcenteredness tends to decline with medical education. Those studies assessed patient centered care but did not specifically examine shared decision making. Patient centered care can be defined as 'the importance of better understanding the experience of illness and of addressing patients' needs'[23], and may or may not include shared decision making. Shared decision making is : "The process by which the optimal decision may be reached for a patient at a fateful health crossroads [...] and involves, at minimum, a clinician and the patient, although other members of the health care team or friends and family members may be invited to participate.[23]" Research focusing on patient centered care suggests that the more experienced medical students become, the less patient-centered they are.[24–29]

We are very aware of the literature surrounding patient centered care, as well as the Patient and Practitioner Orientation Scale (PPOS), and have referenced and carefully reviewed the following studies:

• Tsimtsiou Z, Kirana P-S, Hatzichristou D. Determinants of patients' attitudes toward patient-centered care: A cross-sectional study in Greece. Patient Educ Couns 2014;97:391–5. doi:10.1016/j.pec.2014.08.007

• Hook KM, Pfeiffer CA. Impact of a new curriculum on medical students' interpersonal and interviewing skills. Med Educ 2007;41:154–9. doi:10.1111/j.1365-2929.2006.02680.x

• Pfeiffer C, Madray H, Ardolino A, et al. The rise and fall of students' skill in obtaining a medical history. Med Educ 1998;32:283–8. doi:10.1046/j.1365-2923.1998.00222.x

• Shankar PR, Dubey AK, Subish P, et al. Attitudes of first-year medical students towards the doctor patient relationship. J Nepal Med Assoc 2006;45.

• Shankar PR, Dubey AK, Subish P, et al. Student attitudes towards the doctor-patient relationship in a medical college in western Nepal. Med Teach 2006;28:199.

• Tsimtsiou Z, Kerasidou O, Efstathiou N, et al. Medical students' attitudes toward patient-centred care: a longitudinal survey. Med Educ 2007;41:146–53. doi:10.1111/j.1365-2929.2006.02668.x The PPOS covers patient centered care but does not specifically focuses on decision making and

shared decision making. We have reviewed all 18 items carefully while critically appraising the above literature.

3. The reason I have answered no to study design is largely due a lack of clarity on your understanding of what constitutes shared decision making. Without this I was unable to make a judgment about whether the survey can answer the research question. I think it is important for the reader to understand the reasons for choosing the different clinical scenarios.

We thank the reviewer for their attention to detail. We hope that this concern is addressed by the response to comment 2 and revisions made to the manuscript on page 4.

4. Are you expecting students to answer differently based on whether the patient requires an urgent intervention or optional screening? Personally, as a practicing general practitioner, who is an advocate of shared decision making, I feel there would be a differential response. Perhaps something on your understanding of shared decision making (as this has been contested), would be useful.

We agree with the reviewer and hypothesize that students will answer differently depending on the scenario presented (urgent intervention or optional screening). Given that the present manuscript is a protocol that does not yet report data related to those scenarios, we feel that it is not appropriate to describe the type of answers that we expect to collect.

5. You write that "The information sheet intentionally does not mention SDM, but uses the term "health communication" to reduce potential respondent and desirability biases." However you do use the term "shared decision making" in your survey questions and I am not clear why. Thank you.

We apologize if this sentence was unclear. We used the term health communication in the information sheet in order to minimize the risk of desirability bias in the first question about healthcare decision making and in the scenarios. This was suggested by students who took part in our focus groups and tested the first version of the survey. For subsequent questions about knowledge of shared decision making and attitudes towards shared decision making, we had no choice but to use the term shared decision making. We have revised the following sentences appearing on page 8 of the revised manuscript:

The information sheet intentionally does not mention SDM, but uses the term "health communication" to reduce potential respondent and desirability biases on the first question about healthcare decision making and scenarios. For questions assessing knowledge and attitudes towards SDM, we were forced to use the term SDM.

Reviewer 2

Reviewer Name: Fania R. Gärtner Institution and Country: Leiden University Medical Center, Department of Medical Decision Making, the Netherlands Please state any competing interests:None declared

1. Is the study design appropriate to answer the research question?

No, based on the collected data it is not clear to me how to realize objective 3.: "Determine when and how to best deliver SDM training to medical students." On page 4 you state the objectives and also the aims of the Phase 2 data collection. It is not clear to me how these two exactly relate to each other. Also, it is not clear to me how the two phases relate to each other, are there different research aims for each phase or are they assumed to be analysed together to answer the same research questions?

We thank the reviewer for seeking this clarification. Objective 3 will primarily be realized using data from phase

2. We have revised the objectives listed on page 4 to specify which data collection phase each objective corresponds to.

Consequently, our objectives are to:

 Investigate medical students' knowledge of and attitudes towards SDM across the medical curriculum in four countries, as well as their preferred consultation style (data collected in phase 1);
Investigate the factors that may influence medical students' knowledge of and attitudes towards SDM (data collected in phases 1 and 2);

3) Determine when and how to best deliver SDM training to medical students (data collected in phase 2).

This revision should also address the reviewer's concern about relationship between data collected in phases 1 and 2. To this effect, we have also revised the section describing our research questions and specified which phase of data collection each research question focuses on (see page 6). The following research questions will guide the data collected in phases 1 and 2 of the study:

1) What are medical students' knowledge of and attitudes towards SDM across the medical curriculum? (data primarily collected in phase 1)

2) Do knowledge of and attitudes towards SDM change with medical education? (data primarily collected in phase 1)

3) What are the potential factors that influence SDM during medical education? (data collected in phases 1 and 2)

4) How and when should SDM training be delivered during medical education? (data primarily collected in phase 2)

2. On page 6 you present the research questions. Based on the description of your survey I understand that you measure the following: demographic characteristics, attitude, knowledge, previous training, and based on the scenario's the availability of role models/ their view on the current clinical practice at their institution (If I understood it well) and their behavioural intention. I don't see how you will answer research question 3 and 4 with this data.

As outlined above and in the manuscript, research question 3 (What are the potential factors that influence SDM during medical education?) will be answered using data collected in phases 1 and 2. We will use data collected in phase 1 to do an analysis of covariance (ANCOVA) to evaluate the influence of specific factors such as country, demographics, education level, and previous training on knowledge and attitudes about SDM. This information is provided on page 9 of the manuscript in the analysis section. For phase 2, we will use data collected in the semi-structured interviews with a purposive sample of phase 1 participants and with a convenience sample of curriculum experts. As outlined on page 4 of the revised manuscript, the interviews will help us understand other factors that influence SDM, which may act as barriers and facilitators to teaching SDM in the medical curriculum. In addition, we have included the draft interview guide for phase 2 in Appendix B of the revised manuscript.

3. Are the methods described sufficiently to allow the study to be repeated? No, what is the topic list? I do not fully understand your hypothesis about attitude and knowledge level across the curriculum.

We feel that the reviewer's concern has already been addressed in our responses to the above comments. We are not sure what is meant by 'topic list'. As mentioned above, we have added the phase 2 interview guide to the revised manuscript (appendix B) and hope that it addresses the reviewer's concerns.

4. Analyses of qualitative data:

Unfortunately, I cannot judge about whether phase 2 will provide answers on the research questions, because I miss a topic list.

Do you use the quantitative and qualitative data to answer the same research questions or different ones? I do not regard qualitative data suitable for all research questions. And if you use both phased to answer the same questions (I assume that you plan to do that because you use the word triangulation), how to you combine the data?

The reviewer's concerns have been addressed in responses to comments 1 and 2 above and as shown below:

The following research questions will guide the data collected in phases 1 and 2 of the study:

1) What are medical students' knowledge of and attitudes towards SDM across the medical curriculum? (data primarily collected in phase 1)

2) Do knowledge of and attitudes towards SDM change with medical education? (data primarily collected in phase 1)

3) What are the potential factors that influence SDM during medical education? (data collected in phases 1 and 2)

4) How and when should SDM training be delivered during medical education? (data primarily collected in phase 2)

5. Educational systems differ between countries and curricula differ within countries and also within one institution they differ between cohorts. I wonder if these differences allow to analyse all data together?

We agree with the reviewer. As stated in our analysis plan (pages 9 and 10), we will analyze data within country and between countries. Detailed information is provided on page 9 of the revised manuscript (see below). Regarding the question of data analysis per institution, this might be possible in the US, Canada, and The Netherlands (given we have approached four medical schools in each countries) but will not be possible in the UK (given the number of medical schools approached). This will also be dependent on the number of completed surveys collected at each institution. We have revised our analysis section accordingly:

For phase 1, we will use multivariable analysis to assess differences in knowledge and attitudes about SDM across the curriculum, within each country and between countries. Depending on the total number of completed surveys at each institution within each country, we will also attempt to assess differences in knowledge and attitudes about SDM between institutions. This might be possible in the US, Canada and the Netherlands as four large medical schools have been approached but is unlikely to be achieved in the UK given all medical schools in the country have been approached.

In order to make the cross-country comparison equitable and meaningful, and given that undergraduate medical education ranges from four to six years in the included countries, with variants regarding when the same content is taught or learned, we will use the first year and the last year of medical education only. However, the within country analysis will enable us to compare differences

across all years of undergraduate medical education (up to six years), for each participating country. Depending on the total number of completed surveys at each institution within each country, we will also attempt to assess differences in knowledge and attitudes about SDM between institutions. This might be possible in the US, Canada and the Netherlands as four large medical schools have been approached but is unlikely to be achieved in the UK given that all medical schools in the country have been approached.

We will also use an analysis of covariance (ANCOVA) to evaluate the influence of specific factors such as country, demographics, education level, and previous training on knowledge and attitudes about SDM. To account for any changes in course contents over time we will also include survey month as a control covariate in this analysis.

We plan to perform an analysis of the data after six months of online recruitment in each participating country. We hope that the primary findings will be based on this analysis. However, for practical reasons, should recruitment be slower than expected, we will continue data collection to obtain additional observations in which to test the validity of modeling assumptions and possibly obtain more precise inferences.

6. I assume that interesting differences between the curricula/countries might be obscured.

We agree with the reviewer. It is likely that the sample size at each institution will not enable us to see differences between curricula. However, as discussed above, we will attempt to assess those differences. Differences between countries will be captured.

7. And how do you handle the different samples, students and experts in the analyses? Do you look for differences or do you analyse it as one set of data? In case of separate analyses, I wonder if 6-12 interviews is enough?

As described on page 6 of the revised manuscript, we will be collecting data with up to 12 students per country and up to 12 curriculum experts per country. This represents up to 48 students overall and up to 48 curriculum experts overall. Given this is a purposive sample, we are confident that this sample size is sufficient to achieve data saturation and are following the guidance provided by Guest et al. about sample size and data saturation.

8. You do check for data saturation, please state a pre-defined stopping rule?

As stated on page 6, the stopping rule is 'until data saturation is reached'. This is consistent with the referenced study in the paper (Guest et al.) and most literature on qualitative data analysis and sample size determination.

9. Do the results address the research question or objective? See point 3 and 4.

We have addressed this concern in our answers to comments 1, 2, 3 and 4.

10. Are the discussion and conclusions justified by the results. No, in the Discussion you state "We anticipate that the study findings will help determine the optimal content, format and timing of SDM training in medical education curricula worldwide." Based on the description of your study I am not convinced yet that you can realize this. I might miss some crucial information about your data collection that could convince me about that you indeed will realize these aims.

This manuscript is a protocol. We are therefore not presenting any results. It is thus difficult, at this stage, to determine whether the discussions and conclusion are justified by the results. BMJ Open

does not publish discussions in study protocols and has asked that we delete the discussion section (see editorial comment 2). The sentence pasted above does not appear in the manuscript anymore. In the dissemination section, we have used revised wording instead:

Understanding the perspective of the individuals who manage and coordinate medical education will contribute to determining how to increase the usability, acceptability and effectiveness of future SDM training.

12. Are the study limitations discussed adequately? No, a description of limitations is missing.

As discussed in our response to comment 11, this manuscript is a study protocol and BMJ Open has specifically asked to remove the discussion section. There is therefore no limitation section. However, we have added the strengths and limitations of the study on page 2 of the revised manuscript: Strengths and Limitations of the Study

• We will conduct an international web-based, cross-sectional survey of undergraduate medical students following CHERRIES and COREQ guidelines.

• We followed a comprehensive, iterative survey development process that included several pilot phases.

• In order to determine when and how to deliver SDM training to medical students, this study will also include a stakeholder analysis of medical students and curriculum experts.

• Using convenience samples of medical schools in the US, Canada, and the Netherlands may introduce selection biases.

• Completion of the survey in English by Dutch undergraduate medical students may introduce biases and affect our ability to compare those data across participating countries.

13. Major comments

Protocol

What is the value of a paper describing the protocol for such study? I see the value for such papers in case of an RCT. Please convince me about the value of a publisched protocol for your study.

As outlined on the BMJ open webpage, the publication of study protocols 'enables researchers and funding bodies to stay up to date in their fields by providing exposure to research activity that may not otherwise be widely publicized.' These research activities are not limited to randomized controlled trials and apply to surveys and other types of health services research (including reviews of the literature). Publishing protocol improves the standard of medical research and can limit publication bias and enable replication. In the context of the present manuscript, and given limited research in this area, promoting reproducibility of our survey is key so that this research can be conducted in other countries and in other student cohorts (among nursing students and other allied health professions as well as graduate medical students). Publishing protocols can also prevent unnecessary duplication of work and will hopefully enable collaboration.

14. Relevance and validity

The relevance of this study is not very clear to me. How exactly does your study help curriculum developers, teachers and students?

Also, I expect that the evidence gained in this study is valid only shortly, because curricula change constantly.

The relevance and validity of this study is best outlined in the introduction section and in the dissemination section (see below). Further, given this study is not assessing medical school curricula but aiming to determine (research question 4), among other things, how SDM training should be delivered in the medical school curriculum, we struggle to understand how the changing nature of medical school curricula affects the relevance and validity of the study and related findings. Given the rapidly growing interest in SDM, we argue that our study results will help monitor implementation of

large scale SDM initiatives in the health care system, the education component being a crucial one. As outlined in the introduction:

SDM cannot become widespread unless clinicians fully understand the principles and benefits of SDM, are trained in communicating risks, and engaging patients and significant others (caregivers, family) in deciding about their care. Research suggests that implementing SDM successfully in clinical practice will require interventions targeting the clinicians, the patients, and in the best of worlds, both. Effective interventions targeting clinicians include SDM training.[19] SDM training thus needs to be increasingly embedded in continuing medical education. However, there is little evidence as to which strategies are most effective.[20–22] Yet, continuing medical education is the tip of the iceberg. Training medical students in healthcare communication and SDM seems essential in facilitating routine adoption of SDM in the long term. However, to the best of our knowledge, there is no evidence that the principles of SDM are routinely taught in medical school curricula. Research into the knowledge and attitudes of medical students with regard to SDM is scarce. We have searched the literature, and evidence is also lacking as to when and how to teach SDM principles and skills in medical schools.

As outlined in the dissemination section:

This study is the first to measure medical students' knowledge and attitudes about SDM in Englishspeaking countries, where SDM has been actively promoted but where clinician resistance and lack of understanding of SDM tenets and benefits have significantly limited its widespread adoption.

Understanding the factors that may influence knowledge and attitudes towards SDM to make SDM training particularly beneficial in the undergraduate medical curriculum will be invaluable. Understanding the perspective of the individuals who manage and coordinate medical education will contribute to determining how to increase the usability, acceptability and effectiveness of future SDM training.

15. Scenario's

You describe clinical scenario's that you will use in your survey. I have doubts about their suitability. SDM is often described to be especially valuable for preference sensitive decisions, but in my opinion none of the three cases present a typical preference sensitive decisions. Two scenario's describe decisions about screening. Screening decision are often regarded to be suitable for informed decision making, which differs from SDM in some aspects.

As outlined in the revised manuscript, the scenarios were developed with shared decision making experts. People who are international leaders in the area of shared decision making, and co-authors on this paper (Professors France Légaré and Glyn Elwyn), have confirmed the suitability and relevance of the clinical scenarios as preference sensitive decisions. We define preference sensitive decisions as any decision in which one can not guarantee 100% outcome as planned. We cannot agree with the reviewer that screening decisions are not suitable for shared decision making. Both scenarios B and C are preference-sensitive decisions. We have revised the manuscript on page 6 as follows:

The student survey (see appendix A for English and French versions) comprises five sections:

3) Clinical scenarios where each participant has to indicate: a) how they see other clinicians (e.g., attending physicians, residents, interns) make healthcare decisions and b) how the student would react should they face this situation tomorrow (see Table 1). The clinical scenarios were initially drafted by a Dartmouth fourth year medical student (MW). The first iteration was then revised and reworded by five of the authors, all experts in shared decision making, including two clinicians. The clinical scenarios section also includes one question on risk communication.

16. Language

Dutch students will receive the survey in English language and will be interviewed in English. Please explain why you do not make use of a Dutch survey? English language skills are overall good in the Netherlands but still vary between students.

We have considered this question carefully with our study team and have decided not to translate the questionnaire into Dutch in order to avoid biases that the translation may introduce, and affect the validity of the findings and our ability to compare the findings between countries. Were also faced practical constraints (resources and student volume). We considered translating the survey into Dutch and analyzing and presenting the Dutch data separately from the English data. However, we were concerned about our inability to obtain a large enough Dutch student sample to enable publication of the Dutch data on its own, as well as lose the between-country comparison. Our Dutch collaborator (Dr Johanna Aarts) felt that we would not manage to collect enough data to publish the Dutch data separately, or run both the Dutch and English survey together. In addition, we did not have the necessary resources (available bilingual researchers) to translate the survey in both languages (as was the case for Canada). We will make sure to describe this issue as a limitation in the manuscript reporting the survey findings, and have also added this limitation in the summary that appears on page 2 of the revised manuscript:

Strenghts and Limitations of the Study

• We will conduct an international web-based, cross-sectional survey of undergraduate medical students following CHERRIES and COREQ guidelines.

• We followed a comprehensive, iterative survey development process that included several pilot phases.

• In order to determine when and how to deliver SDM training to medical students, this study will also include a stakeholder analysis of medical students and curriculum experts.

• Using convenience samples of medical schools in the US, Canada, and the Netherlands may introduce selection biases.

• Completion of the survey in English by Dutch undergraduate medical students may introduce biases and affect our ability to compare those data across participating countries.

17. You have pilot tested the survey in Dutch students, but not in year 1 and 2. But these student's will be included in your sample, right? So why no pilot for these first two years for which English language skills can be expected to be lower. I would prefer collecting data in the official language of the school/country. If not possible I would add the response option `I do not understand the question` to prevent any systematic error in the responses of the Dutch students, especially you have set up the survey with forced responses.

For the survey you provide a glossary of terms for non-native English speakers, which might help to overcome language barrier.

Due to practical constraints, we were only able to pilot the study with students in year 3 to 6. The purpose of this pilot was to test the Dutch students' ability to answer the survey in English. We were reassured by our pilot study findings, as outlined in the manuscript, and were confident that English language proficiency was unlikely to differ between students in years 1 and 3. We felt that the use of a glossary would address the proficiency issues that may limit some students' ability to answer all questions in the survey in English. This issue will also be discussed in the limitation section of the survey findings manuscript.

18. For the survey you provide a glossary of terms for non-native English speakers, which might help to overcome language barrier. However, for the interviews such glossary does not work. Especially for qualitative data nuances and exact expression of respondents' opinions, experiences and feelings are very important, therefore I would recommend to interview the students in their own language.

The reviewer makes a valid point. If possible and practical, we will consider conducting the interviews with Dutch students in their native language.

19. The questionnaire

I don't find Table 2 informative. I am not so much interested in the outline itself but much more in the content, which you present in the appendix.

As the reviewer will most likely agree, it would be too difficult and lengthy to present the entire survey in a table, especially given the skip logic and multiple versions of the survey available. We felt that it was a good compromise to present the survey and outline in Table 2 and provide the full survey in appendices.

20. Please provide a reference for the attitude scale you use and the knowledge of SDM scale. Are these valid scales?

And, please explain how attitude can be measured by the OPTION instrument, I know this scale only in the context of measuring the actual process of SDM during medical encounters.

We had made an effort to describe the development and pilot study of the survey in detail, on pages 6, 7, and 8 of the manuscript. We have chosen to assess attitudes towards SDM using OPTION as a proxy for SDM. We have also added the following paragraph on page 7 to address the above questions:

As far as could be determined from our review of the literature, there are no existing validated scales of students' attitudes towards and knowledge of shared decision making available in English. We therefore developed the items presented in appendix A using published literature, and discussion and consensus between study authors. The validated OPTION instrument was initially designed to assess the extent to which practitioners involve patients in decision making processes. We used some of the OPTION items to assess students' attitudes to SDM as well as published studies about clinicians' attitudes to SDM.

VERSION 2 – REVIEW

REVIEWER	Dr. Fania R. Gärtner Department of Medical Decision Making, Leiden University Medical
	Center, The Netherlands
REVIEW RETURNED	24-Apr-2017

GENERAL COMMENTS	Thank you very much for responding so carefully to all my comments. Your responsed very clear and conving and the changes you made to the paper improved the manscript a lot. I have three small comments, two to clearify my first comments and one suggestion for your paper. It is well written and describes a very relevant research question.
	about the previous comment nr. 3: the term "topic-list"indeed refers to the interview guide in semi-structured interviews.
	about the previous comment nr. 8: An example for a stopping rule for data saturations is "if in three subsequent interviews no new information derives, no further interviews are held."
	about the previous comment nr. 19: I might have been not clear, I would suggest leaving Table 2 out in stead of presenting any more information in it.

VERSION 2 – AUTHOR RESPONSE

Reviewer Name: Dr. Fania R. Gärtner Institution and Country: Leiden University Medical Center, Department of Medical Decision Making, the Netherlands Please state any competing interests: None declared

1) Dear authors, thank you very much for responding so carefully to all my comments. Your responsed very clear and conving and the changes you made to the paper improved the manscript a lot. I have three small comments, two to clearify my first comments and one suggestion for your paper. It is well written and describes a very relevant research question.

We thank the reviewer for the positive feedback and agree that the reviewers' comments have significantly improved the manuscript.

2) about the previous comment nr. 3: the term "topic-list"indeed refers to the interview guide in semistructured interviews.

We thank the reviewer for the clarification.

3) About the previous comment nr. 8: An example for a stopping rule for data saturations is "if in three subsequent interviews no new information derives, no further interviews are held."

We thank the reviewer for providing this helpful information and have added a stopping rule on page 6 of the revised manuscript.

"We will aim to recruit a representative sample of up to 12 students per country (up to 48 in total), or until data saturation is reached.[33] The following stopping rule will used: if no new information emerges after three consecutive interviews, no further interviews will be conducted. Students will be offered a \$10 gift card for their participation in these interviews. We will also contact a convenience sample of curriculum experts in each country and ask them to take part in a telephone semi-structured interview. We will aim to recruit a sample of up to 12 curriculum experts per country (up to 48 in total). The same stopping rule, as mentioned above, will be applied for these interviews. The interview guides have already been drafted but will be revised and finalized building on the answers collected in phase 1 (see draft interview guide in supplementary file)."

4) About the previous comment nr. 19: I might have been not clear, I would suggest leaving Table 2 out in stead of presenting any more information in it.

We feel that Table 2 is an important contribution to the manuscript. It will be helpful to readers as it quickly outlines the content of the survey, particularly for those who won't be reviewing the survey file in supplementary files. We suggest keeping Table 2.