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)	Barriers to shared decision-making with women of reproductive age affected by a chronic inflammatory disease: A mixed-methods needs assessment of dermatologists and rheumatologists
AUTHORS	Murray, Suzanne; Augustyniak, Monica; Murase, Jenny; Fischer, Rebecca; Nelson-Piercy, Catherine; Peniuta, Morgan; Vlaev, Ivo

VERSION 1 – REVIEW

REVIEWER	Haske van Veenendaal ESHPM, Erasmus University
REVIEW RETURNED	24-Sep-2020

GENERAL COMMENTS	The authors performed a mixed-methods behavioral and educational needs assessment, employing semi-structured telephone interviews followed by an online survey. This study gives more insight into the reasons why SDM is not fully integrated in dermatology and rheumatology clinical practice. I have several comments for the authors to consider to potentially improve their manuscript. 1. This study focuses on barriers and challenges. Why no facilitators? What is the difference between a barrier and a challenge? 2. Why not conducting the interviews before the survey, that would have led to more in-depth data 3. I am wondering why the authors did not use a theoretical framework for implementation/change. The theories used for the development of the survey, interview list and the coding tree dates from 20 years ago and does not include a SDM model. The perspectives on education might have changed in between. Could relevant information be missed? 4. Although I think that it would have been wise to include more of the implementation/change, I would advise to uses it for reflection on the results in the discussion. Topic 2 and 3 are mainly about behavior (attitudes, knowledge, skills). Topic 4 is about organizational/contextual barriers. That clinicians are not aware of theories, models and guidelines related to SDM refers to the knowledge an tools they need to have for applying SDM. In implementation theory, these are 4 implementation domains at which barriers and facilitators can be present. 5. The primary subject of the paper is SDM, but it seems that by using more general theories as explained under methods, the results of the study are sometimes less relevant to SDM. I.e. Major topic 1 (Barriers to selecting an optimal treatment choice for women of reproductive age) is not an issue of SDM but a lack of professional knowledge. To make it more SDM specific, one could argue that if there's no professional knowledge/consensus about what the best treatment is, it can be an extra reason to apply SDM. 6. For the definition of shared decision making, authors refer to a

	paper Thomas, A., et al.. I find this definition relatively vague compared to other definitions that better describes what happens in the process of SDM (the role of the clinician and patient) and the result of it (that a decision is made that fits best with personal values). 7. Authors present a study of Matthijssen with an OPTION score of 28%. This is somewhat unusual: The OPTION score is generally presented as an absolute number (from 0-100). 8. I would suggest adding a sentence to introduce the objectives of the study to realize a natural transition from text to bullets. 9. It looks like the results only discuss quantitative results? And the qualitative results as quotes? Why not present a (supplementary) table with the themes/topics that came out of the 48 interviews, the most be more to show about these data? 10. In addition, no patients were involved in the research. Could relevant information be missed? 11. I think the discussion and practice implications would improve if the authors think deeper about the prominent role of professional behavior as barrier to SDM. Authors now only refer to the MAGIC-program, but more multilevel/blended approaches have been reviewed, developed and tested. What does this study add to the current knowledge? And given the results of this study, what parts of a multilevel approach would be most promising to test in this setting? 12. Why is shared decision-making not one of the keywords? 13. I would suggest to add the questionnaire and/or interview list as supplementary file.

REVIEWER	Sarah Goff University of Massachusetts, Amherst, USA
REVIEW RETURNED	23-Oct-2020

GENERAL COMMENTS	General: This mixed-methods study explored rheumatologists' and dermatologists' knowledge skills and attitudes related to shared decision making about treatment of chronic inflammatory diseases (CID) during pregnancy. The topic is important and the methodological approach of the study generally sound, but there were a few questions detailed below. One question relates to potential conflicts of interest. Although the authors list their potential conflicts, this alone doesn't eliminate the potential for conflict. A concrete example that raised a potential concern for me was that on my first read I wondered why only one of a number of drugs that could be used during pregnancy was mentioned in the Introduction. When I got to the COI disclosures and looked up the companies the authors worked for or owned, I saw that this is the drug the authors' company makes. On a quick look, it also appears that there is a bit more in the literature on this and related topics. For example, there was a systematic review on influences on physicians' management of women of childbearing age with CID in 2019 that seems to have been missed Nelson-Piercy et al https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-019-4693-x. Overall the authors should be commended on exploring this important topic. Abstract The abstract is generally clear. If possible, a bit more detail about the design might be helpful. Intro

	1. This is a journal with multi-disciplinary readership so It might be helpful if the Introduction there was a brief description of which specific class(es) of drugs the there is clinical equipoise for, any existing guidelines for treatment during pregnancy, what the risks and benefits of treatments are, what are the alternatives, and the potential benefits and harm of a given disease management decision. This would help frame the decision making scenario. 2. Are there other decisional points related to CID and pregnancy besides medication choices amenable to SDM? 3. Is the last sentence missing something perhaps? Methods 1. The use of mixed methods is a strength. The planned use for the interviews appears to have been primarily to inform survey, which is sound practice for questionnaire design. It also appears the interview data was used to augment questionnaire data findings, which is interesting because one would expect that if the interviews were used to inform the survey, there would be relatively similar concepts. 2. Also a strength that it is an international study. 3. Could the authors say a little more about who is on the existing panels they recruited from? Who organizes the panels? Are panelists paid for being on the panel/paid when they participate in a study or other activity related to the panel? Are there data about how panel members may differ from the overall population of dermatologists and rheumatologists in each country (e.g, greater representation from community or academic, age, gender) 4. Was the German translation back translated to English – this is common when translating questionnaires? 5. Who were the expert interviewers? Co-authors? Professionals – if so what profession? If so, who do they work for? 6. Telephone and videoconference are little different methodologically – how was it decided which was used? 7. Was the interview guide pilot tested? If so with whom and how? 8. Who did the qualitative analysis? Was COREQ used? 9. 92 questions seems a lot for 15 minutes. – about 10 seconds/question Was the questionnaire pilot tested? How? If so, what was the range of time it took to complete. Piloting with the study team only seems inadequate. 10. Who beta tested? Was this only about button function? 11. Validation questions can be a strength but it is a little concerning how many surveys were not included due to validation questions. Can the authors do a sensitivity analysis with all of the questionnaires and with different quartiles of inconsistencies to see if/how this changed results? 12. Were the groupings for skills based on response distributions or

	other? Rationale if not response distributions? 13. It seems the authors' position is that dermatologists and rheumatologists might be expected to have expertise on contraception counseling, prescribing and managing. At least in the U.S., this would be an unusual expectation of a specialist. This is one of the areas where the care coordination aspect of the inquiry seems most important. 14. I am not familiar with the Veritas IRB and am curious if there are data on use of external IRBs that the investigators pay for services. It seems like the relationship could have potential for conflict and perhaps it would help toeelaborated on in the conflict statements. 15. Could the questionnaire and interview guides be shared? Results 1. Are the interview and survey populations unique or were some/all of the interviewees in the survey population too? 2. I think it would be helpful to see the interview and survey demographics separately. 3. Although I understand the authors combined the interview and survey data, since the methods state the interviews were done to inform the survey (a strength in survey development) I was expecting to see the qualitative analysis of the interviews reported separately and how they informed questionnaire development specifically. Since the interviews aimed at informing the survey, it doesn't seem that triangulation of the data is really the right term since one would expect the questionnaire to have reflected the preceding interviews. 4. it wasn't entirely clear to me the distinction between decision making regarding planned and unplanned pregnancy and SDM. Do the authors mean that it was harder for derms and rheums to advise once a pregnancy was already in process? Or not clear on how to advise if the woman was on a potentially teratogenic drug and became pregant? Since unplanned pregnancies are probably more common than planned across the globe, this seems likely to not be a rare thing to encounter clinically. 5. Table 2 and 3 – it is not clear to me what criteria were used to categorize as high, intermediate or low suboptimal knowledge since percentages are given below the table but shading does not seem consistent. I found these tables a bit confusing. Discussion 1. The Discussion is clearly presented and well organized. 2. Including some of the literature on SDM in framing of the study's findings would strengthen the Discussion – there is a substantial literature on barriers across many specialties. 3. Strengths and limitations – were data from a source other than the interviews and survey used? I think I only saw these two sources. If so, this strength might be considered a little strong since it is relatively standard best practice methodology to do at least some formative interviews to inform questionnaire development.

	4. Are there national practice guidelines in any of the countries for management of CID when planning for and during pregnancy? This seems important to refer to if so or note there are not. 5. Aspects of the Discussion made me wonder if the concept the authors are thinking about more to is informed decision-making rather than shared decision making. While related, SDM refers specifically to decisions for which there is clinical equipoise. Informed decision-making usually means that there is an evidence-based recommended approach, patients have access to all pertinent knowledge about the evidence, and decision is of course patient-centered, even if the decision contradicts current evidence base. This links to the limited discussion throughout the paper regarding equipoise for treatment options. 6. Agree with rec. to evaluate women and generalist/OB's perceptions if this has not been done. 7. Additional studies presenting cases may not be as useful as some other approaches– these have been shown to reflect actual practice fairly poorly. Audiotaped visits with post- visit interviews with physician and patient (separately) may be more useful. Thank you for tackling this very important communication and care coordination issue!

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Haske van Veenendaal , Erasmus University Rotterdam

Comments to the Author:
The authors performed a mixed-methods behavioral and educational needs assessment, employing semi-structured telephone interviews followed by an online survey. This study gives more insight into the reasons why SDM is not fully integrated in dermatology and rheumatology clinical practice. I have several comments for the authors to consider to potentially improve their manuscript.

1. This study focuses on barriers and challenges. Why no facilitators? What is the difference between a barrier and a challenge?

Response: Authors thank the reviewer for this question. The study objectives have been clarified. One of them is to inform continuous professional development activities and performance improvement interventions, via the identification of unmet educational needs, in the form of barriers and challenges. According to the *Cambridge* English dictionary, a challenge is a "(difficult job) (the situation of being faced with) that needs great mental or physical effort in order to be done successfully and therefore tests a person's ability" (https://dictionary.cambridge.org/dictionary/english/challenge). In comparison, a barrier is (an obstacle) "something that impedes progress or achievement" (Obstacle | Definition of Obstacle by Merriam-Webster (merriam-webster.com). An example of a challenge, or difficult job, is discussing childbearing aspirations and managing unplanned pregnancies, evidently found via the reporting of sub-optimal skills and confidence carrying out these conversations. An example of a barrier, or obstacle, is the perceived lack of data on the compatibility of available therapies for chronic inflammatory diseases during pregnancy, likely preventing clinicians from acquiring the knowledge, skills, and confidence to discuss options with women of reproductive age. Unveiling barriers and challenges is still relevant to future studies interested in exploring facilitators.

2. Why not conducting the interviews before the survey, that would have led to more in-depth data

Response: Authors thank the reviewer for this question and would like to reaffirm that interviews were indeed conducted before the survey. We agree that the strength of qualitative interviews is to collect rich, contextual information. Authors have reviewed the manuscript believe the sequential nature of the mixed-methods study is clearly articulated.

3. I am wondering why the authors did not use a theoretical framework for implementation/change. The theories used for the development of the survey, interview list and the coding tree dates from 20 years ago and does not include a SDM model. The perspectives on education might have changed in between. Could relevant information be missed?

Response: Authors thank the reviewer for this question. A more complete description of theories and models used under 'Methods', 'Theoretical Framework', including the behaviour change wheel, COM-B model, theoretical domains framework and six additional models and perspectives on shared decision making. The reference by Moore (1998) was kept to inform readers on the discrepancy model, which is still relevant in the context of educational and behavioural needs assessments. The authors also acknowledge in the strengths and weaknesses of the study that other types of barriers and challenges, especially at the systemic level, should be further investigated.

Added references:

- Michie, S., L. Atkins, and R. West, The behaviour change wheel. A guide to designing interventions. 1st ed. Great Britain: Silverback Publishing, 2014: p. 1003-1010.
- West, R. and S. Michie, A brief introduction to the COM-B Model of behaviour and the PRIME Theory of motivation [v1]. Qeios, 2020.
- Atkins, L., et al., A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. Implementation Science, 2017. 12(1): p. 1-18.
- Wexler, R. (2012). Six steps of shared decision making. *Informed Medical Decisions Foundation*: 1-2.
- Graffigna, G. and S. Barello (2018). Spotlight on the Patient Health Engagement model (PHE model): A psychosocial theory to understand people's meaningful engagement in their own health care. *Patient Preference and Adherence, 12*: 1261.
- Rennke, S., et al. (2017). The SDM 3 Circle Model: A Literature Synthesis and Adaptation for Shared Decision Making in the Hospital. *Journal of Hospital Medicine 12*(12): 1001-1008.
- Taube, K.-M. (2016). Patient–doctor relationship in dermatology: From compliance to concordance. *Acta Dermato-venereologica, 96*(217): 25-29.
- Beers, E., et al. (2017). The Role of Patients: Shared Decision-Making. *Otolaryngologic Clinics of North America, 50*(4): 689-708.
- (October 2020). "The SHARE Approach." from https://www.ahrq.gov/health-literacy/professional-training/shared-decision/index.html.

4. Although I think that it would have been wise to include more of the implementation/change, I would advise to uses it for reflection on the results in the discussion. Topic 2 and 3 are mainly about behavior (attitudes, knowledge, skills). Topic 4 is about organizational/contextual barriers. That clinicians are not aware of theories, models and guidelines related to SDM refers to the knowledge an tools they need to have for applying SDM. In implementation theory, these are 4 implementation domains at which barriers and facilitators can be present.

Response: Authors believe the reviewer is referring to the behaviour change wheel and the COM-B model and/or the theoretical domains framework. These models are now referenced to strengthen the rationale as to why knowledge, skills, attitude and confidence were assessed, in addition to interprofessional competencies, given the context and health system in which a clinician practice. We thank the reviewer for this recommendation.

Added references:

- Michie, S., L. Atkins, and R. West, The behaviour change wheel. A guide to designing interventions. 1st ed. Great Britain: Silverback Publishing, 2014: p. 1003-1010.
- West, R. and S. Michie, A brief introduction to the COM-B Model of behaviour and the PRIME Theory of motivation [v1]. Qeios, 2020.
- Atkins, L., et al., A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. Implementation Science, 2017. 12(1): p. 1-18.

5. The primary subject of the paper is SDM, but it seems that by using more general theories as explained under methods, the results of the study are sometimes less relevant to SDM. I.e. Major topic 1 (Barriers to selecting an optimal treatment choice for women of reproductive age) is not an issue of SDM but a lack of professional knowledge. To make it more SDM specific, one could argue that if there's no professional knowledge/consensus about what the best treatment is, it can be an extra reason to apply SDM.

Response: Authors thank the reviewer for this comment and agree that SDM is even more relevant when difficult treatment decisions must be made. Authors strengthened the introduction, with additional references to better explain the interplay between SDM and current limitations of available treatment options for women of reproductive age affected with moderate to severe CID. In addition, the discussion has been revised to better position how acquisition of professional competencies, such as knowledge of treatment compatibility during pregnancy and skills discussing risks and benefits of various treatment options, may improve the implementation of SDM.

6. For the definition of shared decision making, authors refer to a paper Thomas, A., et al.. I find this definition relatively vague compared to other definitions that better describes what happens in the process of SDM (the role of the clinician and patient) and the result of it (that a decision is made that fits best with personal values).

Response: Authors thank the reviewers for this comment. Additional references regarding SDM have been added to the introduction to clarify the role of clinicians and patients in shared decision-making, and that the resulting decision is based on both a patient's interest and values, as well as a clinician's expertise and application of evidence-based medicine.

7. Authors present a study of Matthijssen with an OPTION score of 28%. This is somewhat unusual: The OPTION score is generally presented as an absolute number (from 0-100).

Response: Authors thank the reviewer for pointing this oversight. The score should have been presented in absolute number. This has now been corrected in the manuscript.

8. I would suggest adding a sentence to introduce the objectives of the study to realize a natural transition from text to bullets.

Response: Authors thank the reviewer for this recommendation. A transition sentence has been added between the literature review and the study objectives.

9. It looks like the results only discuss quantitative results? And the qualitative results as quotes? Why not present a (supplementary) table with the themes/topics that came out of the 48 interviews, the most be more to show about these data?

Response: Triangulation (i.e., the comparison of multiple sources of information, methods and perspectives) was a critical component of this study. The results section presents both qualitative and quantitative data supporting the main four findings/themes that emerged from this study: 1) barriers to selecting an optimal treatment choice with women of reproductive age, 2) challenges discussing childbearing aspirations and managing unplanned pregnancy, 3) barriers to integrating SDM into clinical practice, 4) challenges collaborating with obstetrician and/or gynaecologists (OB-GYNs) and general practitioners (GPs). Supplementary file #2 (originally supplementary file #1) lists the major topics that emerged from the 48 interviews. This file was provided during original submission and has not changed.

Reference:

- Turner, S.F., L.B. Cardinal, and R.M. Burton, *Research design for mixed methods: A triangulation-based framework and roadmap.* Organizational Research Methods, 2017. **20**(2): p. 243-267

10. In addition, no patients were involved in the research. Could relevant information be missed?

Response: Indeed, this is a limitation of the study (see discussion, 'strengths and weaknesses of the study'). The authors highly recommend additional studies that assess the patient's perspective on dermatologists and rheumatologists' competencies to engage them in shared decision making, especially when discussing the risks and benefits of available therapies for chronic inflammatory diseases during pregnancy. This recommendation is found in the discussion of the manuscript, 'unanswered questions and future research'.

11. I think the discussion and practice implications would improve if the authors think deeper about the prominent role of professional behaviour as barrier to SDM. Authors now only refer to the MAGIC-program, but more multilevel/blended approaches have been reviewed, developed and tested. What does this study add to the current knowledge? And given the results of this study, what parts of a multilevel approach would be most promising to test in this setting?

Response: Authors thank the reviewer for this recommendation. The discussion on implications for clinicians and policymakers has been modified to better articulate how specific professional competencies in dermatology and rheumatology can impede SDM. The discussion also includes a 2020 systematic review of 36 SDM interventions, which concludes that most effective approaches to promoting SM are yet to be further investigated. The present study underscores specific professional competencies necessary for dermatologists and rheumatologists to be able to involve women of reproductive age in SDM, including the presentation of available treatment options before, during and after pregnancy. The present findings move beyond theoretical models of behaviour change and SDM, and present concrete outcomes that can be both measured and addressed through education and training.

12. Why is shared decision-making not one of the keywords?

Response: This keyword is not available as part of the ScholarOne submission system. We added 'patient-centred medicine' as the primary subject heading and 'communication' as the secondary subject heading.

13. I would suggest to add the questionnaire and/or interview list as supplementary file.

Response: Authors thank the reviewer for this recommendation. A new supplementary file (#1) now provides details on interviews and survey questions used to collect information on results presented in the manuscript.

Reviewer: 2
Dr. Sarah Goff, University of Massachusetts Amherst

Comments to the Author:
General: This mixed-methods study explored rheumatologists' and dermatologists' knowledge skills and attitudes related to shared decision making about treatment of chronic inflammatory diseases (CID) during pregnancy. The topic is important and the methodological approach of the study generally sound, but there were a few questions detailed below. One question relates to potential conflicts of interest. Although the authors list their potential conflicts, this alone doesn't eliminate the potential for conflict.

Response: Authors thank the reviewer for this concern and want to reassure that all conflicts of interest were declared, including previous work with the sponsors of this study. Potential conflict of interest was mitigated by having the study 1) reviewed and approved by an Independent Ethics Review Board, 2) led by a third party independently from the sponsor (co-authors from AXDEV Group), and 3) involve a multidisciplinary group of experts, preventing the views of a single author to be over-represented.

A concrete example that raised a potential concern for me was that on my first read I wondered why only one of a number of drugs that could be used during pregnancy was mentioned in the Introduction. When I got to the COI disclosures and looked up the companies the authors worked for or owned, I saw that this is the drug the authors' company makes.

Response: This study was guided by co-authors knowledge and understanding of the field, including current guidelines which explicitly mention certolizumab pegol as a first-line choice for women of reproductive age affected with chronic inflammatory disease during conception, and continued use during all trimesters of pregnancy. The introduction has now been modified to provide readers with a better understanding of therapeutic agents available, and the complexity associated with each class of agents. Existing guidelines are more explicitly referenced, alongside relevant evidence.

On a quick look, it also appears that there is a bit more in the literature on this and related topics. For example, there was a systematic review on influences on physicians' management of women of childbearing age with CID in 2019 that seems to have been missed Nelson-Piercy et al https://can01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fbmchealthservres.biomedcentral.com%2Farticles%2F10.1186%2Fs12913-019-4693-x&data=04%7C01%7C%7C5e8c79d433d445a7505c08d8b6e41899%7C4d4c46e890154f91872a2051c0052397%7C0%7C0%7C637460438094743855%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=Lg3k05S1BZ1crdK6yADUHgluIT0Pey9Dh05cT3m83RA%3D&reserved=0.

Response: Authors have reviewed the manuscript and added additional references. The systematic review by Nelson-Piercy was already included as part of the discussion 'fit within the literature'. It is now also referenced as part of the introduction. The main findings of these studies are summarized and compared with the present study.

Overall the authors should be commended on exploring this important topic.

Response: Authors thank the reviewer for this comment. We strongly believe that this topic necessitates further attention to ensure women of reproductive age affected with a chronic inflammatory disease are provided the best care possible. We are confident that this study will help inform continuous medical educational activities to better support dermatologists and rheumatologists carry critical discussions with this patient population group.

Abstract

The abstract is generally clear. If possible, a bit more detail about the design might be helpful.

Response: We thank the reviewer for this recommendation; the abstract was slightly modified to include information on the triangulation component of this study.

Intro

1. This is a journal with multi-disciplinary readership so It might be helpful if the Introduction there was a brief description of which specific class(es) of drugs the there is clinical equipoise for, any existing guidelines for treatment during pregnancy, what the risks and benefits of treatments are, what are the alternatives, and the potential benefits and harm of a given disease management decision. This would help frame the decision making scenario.

Response: We thank the reviewer for bringing to our closer attention the multi-disciplinarity of the journal's readership. The introduction has been modified to better frame the decision-making scenario. It also provides further details on available classes of agents, risks and benefits associated with each class and/or relevant agents, considering current evidence and expert recommendations. Guidelines and recommendations by the British Association of Dermatologists, the American College of Rheumatology, and the European League Against Rheumatism (EULAR) are now explicitly stated.

2. Are there other decisional points related to CID and pregnancy besides medication choices amenable to SDM?

Response: Sometimes the question is whether pregnancy is too risky for the woman and should be postponed or if she may have difficulties in caring for the child once it is born. This point has been added to the introduction. Further, the results of this study suggest that dermatologists and rheumatologists may experience challenge raising these points during a discussion with patients, as

64% reported sub-optimal skills closely monitoring changes in pregnancy status or childbearing aspirations (see Table 2, items 9 and 10).

3. Is the last sentence missing something perhaps?

Response: The last sentence has been slightly modified for clarification.

Methods

1. The use of mixed methods is a strength. The planned use for the interviews appears to have been primarily to inform survey, which is sound practice for questionnaire design. It also appears the interview data was used to augment questionnaire data findings, which is interesting because one would expect that if the interviews were used to inform the survey, there would be relatively similar concepts.

Response: Authors thank the reviewer for this comment. The interviews were indeed used to inform the development of survey items, in addition to the literature and expertise of co-authors. Both research tools were designed to answer the main study objectives, and therefore included concepts of risk assessment, treatment, and management before, during and pregnancy, as well as patient engagement in shared decision making and interprofessional communication and collaboration among providers. Some findings emerged stronger in one phase of the study over the other, which is why the comparison of data sources at the end of the study was used for identification of convergent and divergent findings (i.e., the triangulation of findings). This allowed co-authors to identify four major themes that were supported by both qualitative and quantitative findings.

2. Also a strength that it is an international study.

Response: Thank you for the comment. This has been added as strength of the study.

3. Could the authors say a little more about who is on the existing panels they recruited from? Who organizes the panels? Are panelists paid for being on the panel/paid when they participate in a study or other activity related to the panel? Are there data about how panel members may differ from the overall population of dermatologists and rheumatologists in each country (e.g, greater representation from community or academic, age, gender)

Response: Different panels were convened for qualitative and quantitative phases. Panels consisted of health care professionals who voluntarily registered to receive invitations for research studies informing interventions and treatments in health care. Panelists are compensated in various ways, depending on the projects they take part of, but are not compensated for simply being on the panel. In the context of this study, the compensations provided were approved as being fair, equitable and non-coercive by the Independent Ethical Review Board (IRB) who approved the protocol. The specific sample was targeted via purposive sampling criteria to ensure a variety of perspectives were captured, depending on the practice settings (e.g., community-based vs. academic-based), ages and genders of participants. This allowed for sub-group analyses to assess differences in clinical practice gaps, challenges and barriers based on socio-demographic variables. Authors were limited with the word count for this manuscript, and therefore, could not include all of this information. We thank the reviewer for providing the opportunity to clarify.

4. Was the German translation back translated to English – this is common when translating questionnaires?

Response: Although the German interview guide and questionnaires were not translated back to English, they were reviewed by peer translators as an additional quality assurance checkpoint. This information was added to the manuscript.

5. Who were the expert interviewers? Co-authors? Professionals – if so what profession? If so, who do they work for?

Response: The expert interviewers were two of the co-authors (MA and MP, researchers at AXDEV Group), who trained additional professional interviewers during a briefing session to ensure alignment

with the study objectives, main questions, and interviewing techniques. All interviewers have a background in health care. This information was added to the manuscript.

6. Telephone and videoconference are little different methodologically – how was it decided which was used?

Response: All participants received a dial phone number to participate in a conference call (without video) in addition to a link in case the participant preferred joining by computer or phone application. The manuscript was adjusted to specify interviews were conducted over a conference call.

7. Was the interview guide pilot tested? If so with whom and how?

Response: The interview guide was not pilot tested. However it was reviewed by multiple researchers at AXDEV Group (an organization with more than 23 years of experience in this field of research). The semi-structured nature of interviews allowed the interviewers to adjust questions based on the participant's understanding of the subject and communication style. All interviewers were instructed to communicate any issues to the rest of the research team immediately so these could be addressed in future interviews.

8. Who did the qualitative analysis? Was COREQ used?

Response: The coding was completed by MA, SL and SZ (researchers at AXDEV Group). The thematic analysis was completed by MA and MP. The COREQ reporting checklist was not used, as this is a mixed-method research study. Authors followed the following criteria in reporting mixed-methods studies:

- Describe the justification for using a mixed-method approach to the research question (see study design)
- Describe the design in terms of the purpose, priority and sequence of methods (see methods, study design)
- Describe each method in terms of sampling, data collection and analysis (see methods, data collection and data analysis, respectively)
- Describe where integration has occurred, how it has occurred and who has participated in it (see methods, study design and data analysis, describing triangulation)
- Describe any limitation of one method associated with the presence of the other method (see discussion, strengths and weaknesses of the study)
- Describe any insights gained from mixing or integrating methods (see results)

O'cathain, A., et al. (2008). The quality of mixed methods studies in health services research. *Journal of health services research & policy 13*(2): 92-98.

9. 92 questions seems a lot for 15 minutes. – about 10 seconds/question Was the questionnaire pilot tested? How? If so, what was the range of time it took to complete. Piloting with the study team only seems inadequate.

Response: The survey consisted of a maximum of 16 main questions. Most questions asked participant to rate their knowledge, skill, confidence, or agreement with key items. The number of items per questions varied between 4 to 23, adding to a maximum total of 92 items. Based on similar studies conducted by AXDEV Group (see references below), the allocated time and number of questions and related items was deemed sufficient for these types of assessments. The questionnaire was critically reviewed by all co-authors and tested by MA (researcher at AXDEV Group), to ensure the programming of the survey was correct. No issues in navigation occurred and estimated time for completion did not exceed 15 minutes. The survey was then launched within a small sample of potential participants, 63 of which returned completed. No issues were identified; therefore, the survey was deployed among the remaining pool of potential participants. The median time of completion was 12 minutes and 45 seconds.

References:

Murray, S., et al. (2019). Needs and challenges among physicians and researchers in thrombosis and hemostasis: Results from an international study. *Research and Practice in Thrombosis and Haemostasis 3*(4): 626-638.

Murray, S., et al. (2015). International challenges in patient-centred care in fertility clinics offering assisted reproductive technology: providers' gaps and attitudes towards addressing the patients' psychological needs. *Journal of European CME 4*(1): 27578.

10. Who beta tested? Was this only about button function?

Response: The beta-testing was completed by MA (researcher at AXDEV Group), to ensure the programming of the survey was correct. No issues in navigation occurred and estimated time for completion did not exceed 15 minutes.

11. Validation questions can be a strength but it is a little concerning how many surveys were not included due to validation questions. Can the authors do a sensitivity analysis with all of the questionnaires and with different quartiles of inconsistencies to see if/how this changed results?

Response: The authors do not see the need to complete a sensitivity analysis to assess the impact on results for including 'careless respondents', since the criteria for exclusion based on poor attention were determined a priori. An additional reference was added as part of the manuscript with regard to this method. As a note, many of the 'careless respondents' had "repeated or pattern responses" (e.g., 1-1-1-1-1, or 1-2-3-4-5), which would negatively impact the reliability of results.

12. Were the groupings for skills based on response distributions or other? Rationale if not response distributions?

Response: Skill ratings of none, basic and intermediate were grouped as sub-optimal, meanwhile those rated as advanced or expert were grouped as optimal. Authors believe that ratings below 'advanced' are a concern and indication for improvement. This method has been used and published in the past.

References:

Murray, S., et al. (2019). Needs and challenges among physicians and researchers in thrombosis and hemostasis: Results from an international study. *Research and Practice in Thrombosis and Haemostasis 3*(4): 626-638.

Murray, S., et al. (2015). International challenges in patient-centred care in fertility clinics offering assisted reproductive technology: providers' gaps and attitudes towards addressing the patients' psychological needs. *Journal of European CME 4*(1): 27578.

13. It seems the authors' position is that dermatologists and rheumatologists might be expected to have expertise on contraception counseling, prescribing and managing. At least in the U.S., this would be an unusual expectation of a specialist. This is one of the areas where the care coordination aspect of the inquiry seems most important.

Response: We agree with the comment: dermatologists and rheumatologists should not be expected to counsel on contraceptive use at the same level of expertise as obstetricians and/or gynaecologists. This is why participants were asked to rate their level of skill discussing contraceptive methods with patients (in general), "according to what is expected of them in their professional role." Further, participants were asked to select "not relevant" if they believed this skill was not relevant to their current professional role. This information is now found in a new supplementary file (#1), which contains the main questions included in the interview guides and survey.

14. I am not familiar with the Veritas IRB and am curious if there are data on use of external IRBs that the investigators pay for services. It seems like the relationship could have potential for conflict and perhaps it would help toeelaborated on in the conflict statements.

Response: Veritas IRB (https://www.veritasirb.com/index.html) is accredited by Human Research Accreditation Canada (www.hracanada.org) and operates according to international regulations and guidelines such as the World Medical Association Declaration of Helsinki, the US Code of Federal Regulations and the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans. The authors do not have any existing relationships with the IRB reviewer who reviewed and approved this study for meeting best practice in ethical research. Due to word count limitation, this information could not be added.

15. Could the questionnaire and interview guides be shared?

Response: We added a framework of the research tools with examples of questions asked in respective phases of the study (see supplementary file 1).

Results

1. Are the interview and survey populations unique or were some/all of the interviewees in the survey population too?

Response: Participants were recruited from different panels at different points in time for the interviews and survey respectively. If an individual was included in both panels, they might have been able to participate in both phases of the study. This should not impact the quality and or rigour of the study.

2. I think it would be helpful to see the interview and survey demographics separately.

Response: Thank you for the suggestion. Authors have modified Table 1 to present interview and survey demographics separately.

3. Although I understand the authors combined the interview and survey data, since the methods state the interviews were done to inform the survey (a strength in survey development) I was expecting to see the qualitative analysis of the interviews reported separately and how they informed questionnaire development specifically. Since the interviews aimed at informing the survey, it doesn't seem that triangulation of the data is really the right term since one would expect the questionnaire to have reflected the preceding interviews.

Response: Triangulation is a critical component of mixed-methods research studies and is performed at the end of the data collection and analyses phases to compare different sources, methods, and analyses to identify convergent and divergent findings. This helps to minimise bias that comes from single-method, single-perspective studies. Even though the survey was informed by the interviews, researchers left opportunity for potential differences and/or similarities in findings.

4. it wasn't entirely clear to me the distinction between decision making regarding planned and unplanned pregnancy and SDM. Do the authors mean that it was harder for derms and rheums to advise once a pregnancy was already in process? Or not clear on how to advise if the woman was on a potentially teratogenic drug and became pregant? Since unplanned pregnancies are probably more common than planned across the globe, this seems likely to not be a rare thing to encounter clinically.

Response: Authors investigated challenges and barriers to treating and managing women of reproductive age with a CID before, during and after pregnancy. Challenges were found at all stages of the reproductive journey but appeared more prominent during pregnancy and breastfeeding, especially if these were unplanned. The issue of unplanned pregnancy is common and relates to challenges experienced by physicians when engaging women of reproductive age in discussions that relate to their childbearing aspirations and use of contraceptives (before conception). If an unplanned pregnancy occurs, the discussion transitions more into informing the patient about the risks and benefits of adjusting the treatment regimen. This discussion is especially challenging when it involves newly approved drugs that do not have rigorous data on pregnancy outcomes.

5. Table 2 and 3 – it is not clear to me what criteria were used to categorize as high, intermediate or low suboptimal knowledge since percentages are given below the table but shading does not seem consistent. I found these tables a bit confusing.

Response: Authors have reviewed the shading based on the criteria described at the bottom of each table, and do not see any inconsistencies. To avoid confusion, authors have decided to remove the shading.

Discussion

1. The Discussion is clearly presented and well organized.

Response: Thank you.

2. Including some of the literature on SDM in framing of the study's findings would strengthen the Discussion – there is a substantial literature on barriers across many specialties.

Response: Additional references have been added, including an evaluation of the MAGIC program and a systematic review of 36 SDM interventions.

3. Strengths and limitations – were data from a source other than the interviews and survey used? I think I only saw these two sources. If so, this strength might be considered a little strong since it is relatively standard best practice methodology to do at least some formative interviews to inform questionnaire development.

Response: An initial review of the literature was used and discussed among the multidisciplinary group of authors to inform areas of exploration. Since authors were limited with the word count of the manuscript, it was decided that these details would be omitted, especially that literature reviews are the basic step (and best practice) for all research initiatives.

4. Are there national practice guidelines in any of the countries for management of CID when planning for and during pregnancy? This seems important to refer to if so or note there are not.

Response: While several recommendations exist in various countries (USA: Sammaritano, UK: Flint, EULAR: Skorpen), and are cited as part of the introduction. Recommendation such as these are only as strong and valuable as the data form which they draw and to which they refer. These recommendations are frequently not implemented because many physicians feel insecure about the rigour of the evidence itself, as identified by a systematic review by Nelson Catherine Piercy.

Reference:

- Nelson-Piercy, C., et al., *What factors could influence physicians' management of women of childbearing age with chronic inflammatory disease? A systematic review of behavioural determinants of clinical inertia.* BMC health services research, 2019. **19**(1): p. 1-9.

5. Aspects of the Discussion made me wonder if the concept the authors are thinking about more to is informed decision-making rather than shared decision making. While related, SDM refers specifically to decisions for which there is clinical equipoise. Informed decision-making usually means that there is an evidence-based recommended approach, patients have access to all pertinent knowledge about the evidence, and decision is of course patient-centered, even if the decision contradicts current evidence base. This links to the limited discussion throughout the paper regarding equipoise for treatment options.

Response: The authors ensured the concept of shared decision making was central to the study objectives and design of data collection tools and data analysis. Multiple models and perspectives of shared decision making were reviewed in this process. These are now cited as part of the introduction, and discussion, especially as they relates to limited clinical evidence on the risks and benefits of available treatments for women of reproductive age before, during and after pregnancy.

6. Agree with rec. to evaluate women and generalist/OB's perceptions if this has not been done.

Response: Thank you. We are still not aware of such studies to date. This recommendation has not changed.

7. Additional studies presenting cases may not be as useful as some other approaches– these have been shown to reflect actual practice fairly poorly. Audiotaped visits with post- visit interviews with physician and patient (separately) may be more useful.

Response: Thank you for the recommendation. We agree with the reviewer that audiotaped visits could offer a more objective way of assessing patient-provider interactions and identifying existing challenges and barriers from and educational and behavioural perspective. We have added this to the discussion.

Thank you for tackling this very important communication and care coordination issue!

Response: Thank you for taking the time to critically review and provide feedback on this important study.

VERSION 2 – REVIEW

REVIEWER	Haske van Veenendaal ESHPM, Erasmus University
REVIEW RETURNED	22-Apr-2021

GENERAL COMMENTS	This is a revised manuscript of a mixed-methods study, employing semi-structured interviews followed by an online survey. This study gives more insight into the reasons why SDM is not fully integrated into dermatology and rheumatology clinical practice. The authors have been explaining in detail how they have been processing the feedback of the reviewers including mine. In general, I think they have done a good job. I have some minor comments for the authors to help improve their manuscript: - please perform a grammar check; - please define when a participant is a 'careless responder'. How many were careless? - table 2, item 7, also seems to cross the significance level? - table 3. Does it seem that respondents underestimate the extent to which they think that they facilitate SDM? Maybe interesting for the discussion? - More relevant theory is used for the concept of implementation, such as Scholl, and for SDM. There's a lot to choose from. Nevertheless, I wonder if the authors have looked at standard literature such as the SDM-model of Elwyn?

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Dr. Haske van Veenendaal, Erasmus University Rotterdam

Comments to the Author:

This is a revised manuscript of a mixed-methods study, employing semi-structured interviews followed by an online survey. This study gives more insight into the reasons why SDM is not fully integrated into dermatology and rheumatology clinical practice. The authors have been explaining in detail how they have been processing the feedback of the reviewers including mine. In general, I think they have done a good job.

Response: The authors thank the reviewer for this comment.

I have some minor comments for the authors to help improve their manuscript:

- please perform a grammar check;

Response: The paper was reviewed for grammar errors. Thank you.

- please define when a participant is a 'careless responder'. How many were careless?

Response: A careless responder was defined as a participant who responded without regard to item content. This definition is now included in the text of the manuscript (second to last sentence of the section “Quantitative phase: Online Survey.” Original figure 1, demonstrates that 107/447 respondents were excluded due to “low quality response”. For consistency, the revised figure 1 states “careless responding” and the new description now includes detail as to how participants were considered “careless”. Further, an additional line is reported under “strengths and weaknesses of the study” to tentatively explain this number and that survey fatigue may have played a role.

- table 2, item 7, also seems to cross the significance level?

Response: The p-value of item 7 “discussing contraceptive methods with patients”, in table 2, indeed crosses the significance level. Authors have already indicated this with an asterisk. Therefore, no change was required.

- table 3. Does it seem that respondents underestimate the extent to which they think that they facilitate SDM? Maybe interesting for the discussion?

Response: We thank the reviewer for this suggestion. Two key findings (items 5 and 6) are reported in table 3, underscoring the suboptimal integration of SDM into clinical practice. Items support the difficulty of supporting reflective motivation when thinking about engaging patients in SDM and the presence of environmental norms or constraint (i.e., time pressure). These have been added to the results text and are now discussed under “fit within the literature” as a potential indication that some participants overestimate their implementation of SDM as indicated by other studies. This corroborates with the present study finding that only 25% of participants report often forgetting to engage their patients in SDM or making treatment decisions without the input of their patient.

- More relevant theory is used for the concept of implementation, such as Scholl, and for SDM. There's a lot to choose from. Nevertheless, I wonder if the authors have looked at standard literature such as the SDM-model of Elwyn?

Response: The last sentence under section “Theoretical Framework” mentions “various models and perspectives of SDM were used in the collection and analysis of data”. Seven references are provided, including Elwyn’s SDM model (see reference #20).