PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Determinants of women's preferences for surgical versus conservative management for pelvic organ prolapse: a survey-based study from Italy
AUTHORS	Ferrari, Amerigo; Bellè, Nicola; Giannini, Andrea; Simoncini, Tommaso; Vainieri, Milena

VERSION 1 – REVIEW

REVIEWER	Zhuo, Yueran Mississippi State University, Marketing
REVIEW RETURNED	09-Feb-2024

GENERAL COMMENTS	This paper presents a nationwide prospective study aimed at identifying the determinants of women's choice between surgical and conservative management of Pelvic Organ Prolapse (POP) in Italy. The study utilizes a Qualtrics questionnaire distributed via social media to Italian women over 18, exploring their likelihood of opting for surgery under various scenarios. The research addresses an important question within the field of gynecology, particularly regarding patient preferences and the factors influencing decision-making in POP treatment. The topic is very relavent to practice and the methodology is novel. I have the following concerns and recommendations to further improve this manuscript. 1. The choice to survey a broad population of women, predominantly healthy and potentially without direct experience or risk of POP, may not yield insights reflective of actual patient preferences facing POP management decisions.I recommend the authors to refocus the survey to target women with experiences or characteristics more closely aligning with those at risk of POP, such as women with multiple childbirth experiences. This approach could provide a more accurate reflection of the preferences and considerations of potential POP patients. 2. The absence of control questions within the survey to assess the reliability of respondents' answers limits the ability to gauge the validity of the responses in reflecting genuine preferences regarding POP treatment. Incorporate control questions related to medical decisions with well-understood patient preferences, such as epidural usage during childbirth. This would allow for a comparison of survey responses with known data, offering a measure of the survey's reliability. 3. The study lacks direct input from actual POP patients, both those facing treatment decisions and those who have undergone treatments, to compare and validate the surveyed preferences of

	the broader population. I with the authors could include a comparative analysis with actual POP patients regarding their treatment preferences and post-treatment reflections. This could significantly enhance the study's relevance and applicability to clinical practice. 4. The survey does not adequately distinguish between different surgical options, particularly between obliterative and reconstructive surgeries, which have significant implications for patient quality of life and satisfaction. Please consider add questions to the survey that specifically address the preferences for obliterative versus reconstructive surgery, considering their impact on sexual function and patient well-being. While the research question is of significant interest and relevance, the current study design and methodology present limitations that may impact the applicability and reliability of the findings. A more targeted approach to participant selection and survey design is recommended to ensure the study's outcomes can effectively inform clinical practice and health policy. Given the concerns outlined regarding the study's target population, methodology, and differentiation of treatment options, I recommend a very major revision of the manuscript. The suggested modifications aim to enhance the study's validity, relevance, and contribution to the field.

REVIEWER	Collins, Sarah A The University of Chicago Hospital, Obstetrics & Gynecology
REVIEW RETURNED	18-Mar-2024

GENERAL COMMENTS	I applaud the authors for this innovative way of accessing patient values and preferences around prolapse surgery and exploring some of the factors that should be considered in shared decision making around treatment for pelvic organ prolapse. I have several comments that, if addressed, could improve the manuscript: 1) In the abstract, leading / opinionated statements do not belong in the background / objective section. Additionally, the statement "variation should be driven just by patients' preferences and needs" could be interpreted as violating the major principles of shared decision-making. Therefore, I would omit this statement. 2) The definition of prolapse offered in line 67 of the introduction should include a formally referenced citation, probably the IUC POP Clinical Definition of Prolapse (IUGA). 3) The methods and results sections could benefit from a close English language edit to ensure accurate representation of the study. 4) In the primary outcome section of the results, why do the authors use "In contrast" to introduce factors with odds ratios over 2.8? To me, it doesn't seem as if there is a contrast between odds ratios of 2.9 and 2.8, and it might be appropriate to group these together, unless there is a different point the authors are trying to make that I don't understand. 5) To clarify, does the 61.5% median for overall preference for surgery include analysis of all 11 questions, just the reference question, or just the 10 factor questions? 6) Do the "previous studies" mentioned on line 231 involve survey-based research, like this study, or are they clinical studies in which surgery was actually selected by fewer patients than this study would indicate?

	7) Please elaborate on the claim that the clinical scenario was "not easy to understand" (line 282). Why would this be? Was the wording poor or the language too advanced? 8) In the authors' conclusions in both the Abstract and the paper's Discussion section, the authors assert that the study results reveal a relationship between "unwarranted variation" and determinants of patient choice around prolapse treatment. However, if "unwarranted variation" involves factors NOT related to patients, then I don't think this is an accurate statement. This study seems to identify factors that influence patient preference around surgical versus non-surgical management and does NOT shed any light on sources of unwarranted variation, such as surgeon skill / preference, hospital resources, and regional medical infrastructure. Please clarify this or resolve this inconsistency.

REVIEWER	Brabers, Anne NIVEL Netherlands Institute for Health Services Research
REVIEW RETURNED	19-Mar-2024

GENERAL COMMENTS	Dear authors, Thank you for giving me the opportunity to review this article about practice variation, a topic in which I am highly interested. Hereafter, I will provide my feedback on the manuscript. One of my main concerns is that you performed your study among women of the general population, most of them not suffering from POP. I am in doubt whether research among this group gives a good insight in patient related factors influencing the preference for surgery or not. Another main concern is whether you really get insight in individual determinants. When I think about individual determinants, I think about for example attitude, beliefs, preferences, involvement, information (see e.g. the Theory of Reasoned Action from Fishbein & Ajzen (2010)). But also factors like health literacy, and social norms (e.g. whether women in the social environment of the woman have had a surgery for POP). The last main concern is the link between the outcomes of the study and the conclusion that the variation is strictly related to physician advice. I doubt whether information on clincial outcomes depends fully on the phsycian, is there no general information available about clinical outcomes available (e.g. in a decision aid), or on the internet? Based on the fact that advice from a gynaecologist increases the choice for surgery, the authors conclude that variation is mainly due to the physician. Althoug, I understand and support this (based on other literature), I think ,this conclusion could be better substantiated. For example, the results also show that 80% opt for surgery when there is a hospital nearby, and that is not a factor related to the physicians attitude. Furthermore, I expected that the analyses were done by region. If preferences of women do not differ between regions, the differences in regions as shown in Figure 1 (but also in whole Italy) are not related to the preferences of the women, but to other factors. Is it possible to do the analyses per region? In the abstract the authors write: 'and partly capture the individual determinants of unwarranted variation'. In the discussion the authors state that it is strictly the physician that causes variation. I should bring this more in line.

	In the introduction, line 74, the authors talk about the benefits of conservative treatment. But they are not explained, what are the benefits? On line 82, the authors already argue that practice patterns mainly depend on the surgical training. If this is the case, and I think this is among others the case, what is then the added value of looking into preferences of women to explain variation. I think the authors can argue a bit more in the introduction why looking into the role of women is interested to explain practice variation. Regarding the methods, an important point is the selection of respondents. Why did the authors not invite only women who have POP? It appears that they even have tried to reach the target group. Line 129 in the methods, wat is grade 3 of POP The authors included advantages and disadvantages in the questionnaire. On what kind of information was this based? I think this is important as the way how information is framed can also influence the choice. Why is a scale from 1-99% chosen, and not 0-100? And on line 166 it is stated that it is converted in a 0-1 scale, how was this done? I do not understand whether everyone only saw one of the 11 questions, or saw all questions. On line 132 it is stated that each women received only one question out of 11. While on line 142 it is stated that women further received the other 10 questions in a nonrandomized manner. In table 2 it appears that only 1 question is filled out by each women. If that is the case, I think 17-23 women per factor is not that much. Regarding the chosen factors, I understand that less information was available about the topic, but why did the authors search for individual factors that are related to patient choices, and adapt that to the situation? Like attitudes. For now it appear a bit a random choice of factors, and what is for example the difference between advice from a GP and advice from two physicians. Furthermore, the five gynaecologists that are consulted where all from one university, they probably think and do the same as physicians adapt to each other. Why were not physicians invited from different hospitals/regions? The questionnaire was piloted, in line 161-162 it is stated whether women in the pilot had other factors or determinants that were not considered by the researchers. Could the authors give some insight in whether they adapt something in the questionnare based on the pilot. In the discussion, line 232/233 the authors report that other studies have lower choice rates for POP surgery. Could they explain these differences with their own study? In line 234 the authors use the term 'received no input'. I think that is not the right phrase, as also those women receive input, namely from the scenario. So I think, this has to be formulated in another way. On line 303/305 the authors argue that guidelines can help to reduce practice variation. From the literature, we know that guidelines not reduce variation. See for example: Jong, J.D. de, Groenewegen, P.P., Spreeuwenberg, P., Schellevis, F., Westert,

	G.P. Do guidelines create uniformity in medical practice? Social Science & Medicine: 2010, 70(2), 209-216. Could the authors explain why they think that guidelines will work in this situation? Furthermore, in the literature it is suggested that SDM (with the help of decision aids) can reduce practice variation. E.g. O'Connor AM, Llewellyn-Thomas HA, Flood AB. Modifying unwarranted variations in health care: shared decision making using patient decision aids. Health Affairs 2004; Var63-72 / Greer AL, Goodwin JS, Freeman JL, Wu ZH. Bringing the patient back in. Guidelines, practice variations, and the social context of medical practice. International Journal of Technology Assessment in Health Care 2002; 18(4):747-61 / O'Connor AM, Wennberg JE, Legare F, Llewellyn-Thomas HA, Moulton BW, Sepucha KR, et al. Toward the 'tipping point': decision aids and informed patient choice. Health Affairs 2007; 26(3):716-25 / Stiggelbout AM, Van der Weijden T, De Wit MP, Frosch D, Legare F, Montori VM, et al. Shared decision making: really putting patients at the centre of healthcare. The BMJ 2012; 344:e256. An empirical study also confirms this: Brabers AEM, van Dijk L, Groenewegen PP, et alDoes a strategy to promote shared decision-making reduce medical practice variation in the choice of either single or double embryo transfer after in vitro fertilisation? A secondary analysis of a randomised controlled trialBMJ Open 2016;6:e010894. doi: 10.1136/bmjopen-2015-010894. I think the authors have to include some literature about this topic. They mention in line 310 a well informed decision-making process, but I think they have to refer to the literature about SDM, and say something about that.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Ms. Yueran Zhuo, Mississippi State University
This paper presents a nationwide prospective study aimed at identifying the determinants of women's choice between surgical and conservative management of Pelvic Organ Prolapse (POP) in Italy. The study utilizes a Qualtrics questionnaire distributed via social media to Italian women over 18, exploring their likelihood of opting for surgery under various scenarios. The research addresses an important question within the field of gynecology, particularly regarding patient preferences and the factors influencing decision-making in POP treatment. The topic is very relavent to practice and the methodology is novel. I have the following concerns and recommendations to further improve this manuscript.

Comment 1. The choice to survey a broad population of women, predominantly healthy and potentially without direct experience or risk of POP, may not yield insights reflective of actual patient preferences facing POP management decisions.I recommend the authors to refocus the survey to target women with experiences or characteristics more closely aligning with those at risk of POP, such as women with multiple childbirth experiences. This approach could provide a more accurate reflection of the preferences and considerations of potential POP patients.
Response. Thank you for this comment that may enable us to better explain the rationale of our study. It is true that we could have included only women with higher risk of POP or who suffered from POP in the study, but we decided to include women from the general population on purpose. In fact, we believe that this allowed us to identify the determinants of choice at an early stage, before the condition occurred and the women had visited a health professional. By that time, the woman may have already received input that in some way may influence her choice of treatment. Since our main

purpose was to investigate the effect of individual determinants of choice on practice variation, we inevitably had to choose this type of population to answer the research question. In addition, only 10% of respondents said they did not know pelvic organ prolapse at all. Also, the pilot phase of the study demonstrated the easy comprehensibility of the clinical scenario. We have tried to better explain the rationale behind our choice in the section "Limitations and strengths." Including women who suffer from POP or are at greater risk of developing POP would mean redesigning and conducting the study all over again, so we defer the choice to the editors.

Comment 2. The absence of control questions within the survey to assess the reliability of respondents' answers limits the ability to gauge the validity of the responses in reflecting genuine preferences regarding POP treatment. Incorporate control questions related to medical decisions with well-understood patient preferences, such as epidural usage during childbirth. This would allow for a comparison of survey responses with known data, offering a measure of the survey's reliability.
Response. Thank you for the suggestion. We re-contacted the women through Qualtrics and administered the control question to them (“Would you use epidural analgesia during childbirth?”). We have edited the method and result sections accordingly.

Comment 3. The study lacks direct input from actual POP patients, both those facing treatment decisions and those who have undergone treatments, to compare and validate the surveyed preferences of the broader population. I with the authors could include a comparative analysis with actual POP patients regarding their treatment preferences and post-treatment reflections. This could significantly enhance the study's relevance and applicability to clinical practice.
Response. Thank you for this comment, which is very similar to Comment 1. We hope we have properly explained in our response to that previous comment the rationale for our choice to enroll women from the general population. Although we do not believe that enrolling an additional study population (women suffering from or at higher risk risk of POP) as a comparison group would help us answer the research question, we are willing to do so if the reviewers/editors deem it necessary.

Comment 4. The survey does not adequately distinguish between different surgical options, particularly between obliterative and reconstructive surgeries, which have significant implications for patient quality of life and satisfaction. Please consider add questions to the survey that specifically address the preferences for obliterative versus reconstructive surgery, considering their impact on sexual function and patient well-being.
Response. Thank you for this comment. Actually, we think we partly addressed this point through this question: "How likely would you choose to have surgery if you knew for sure that you would have no scars on your abdomen after the operation?". It is true that there is no clear distinction between reconstructive and obliterative surgery. However, we think this specification may go beyond the specific purpose of the study, partly because it is perhaps too technical a distinction. In the end, the result that the woman sees is the presence or absence of abdominal scarring. So, we decided to focus more on vaginal vs. abdominal surgery rather than on vaginal reconstructive vs. obliterative surgery, believing that women were more interested in the cosmetic/esthetic outcomes rather than the specific surgical technique performed. We have added this limitation to the “Limitations and strengths” section.

While the research question is of significant interest and relevance, the current study design and methodology present limitations that may impact the applicability and reliability of the findings. A more targeted approach to participant selection and survey design is recommended to ensure the study's outcomes can effectively inform clinical practice and health policy. Given the concerns outlined regarding the study's target population, methodology, and differentiation of treatment options, I recommend a very major revision of the manuscript. The suggested modifications aim to enhance the study's validity, relevance, and contribution to the field.

Reviewer: 2
Dr. Sarah A Collins, The University of Chicago Hospital
I applaud the authors for this innovative way of accessing patient values and preferences around prolapse surgery and exploring some of the factors that should be considered in shared decision making around treatment for pelvic organ prolapse. I have several comments that, if addressed, could improve the manuscript.

Comment 1. In the abstract, leading / opinionated statements do not belong in the background / objective section. Additionally, the statement "variation should be driven just by patients' preferences and needs" could be interpreted as violating the major principles of shared decision-making. Therefore, I would omit this statement.
Response. Thank you for the suggestion. We have removed the sentence "variation should be driven just by patients' preferences and needs". We have also edited the background / objective section of the Abstract by smoothing subjectivity as much as possible.

Comment 2. The definition of prolapse offered in line 67 of the introduction should include a formally referenced citation, probably the IUC POP Clinical Definition of Prolapse (IUGA).
Response. Thank you for the comment. We have added the suggested citation.

Comment 3. The methods and results sections could benefit from a close English language edit to ensure accurate representation of the study.
Response. We modified the language to our best. Is there any grammatical error, typo or unclear passage that the Reviewer would like to point out? Can the Reviewer provide any more specific suggestions or highlight where the text is less clear?

Comment 4. In the primary outcome section of the results, why do the authors use "In contrast" to introduce factors with odds ratios over 2.8? To me, it doesn't seem as if there is a contrast between odds ratios of 2.9 and 2.8, and it might be appropriate to group these together, unless there is a different point the authors are trying to make that I don't understand.
Response. We have replaced “In contrast” with “…, while it was…”. We hope this sounds better now.

Comment 5. To clarify, does the 61.5% median for overall preference for surgery include analysis of all 11 questions, just the reference question, or just the 10 factor questions?
Response. This value (61.5%) represents the median of the percentages given as answers to the first randomized question in the entire study population, thus including all 11 questions. We have modified this sentence of the result section as follows: “In the overall population (n=222), the median probability of choosing surgical treatment was 61.5%, regardless of which of the 11 questions was asked first”.

Comment 6. Do the "previous studies" mentioned on line 231 involve survey-based research, like this study, or are they clinical studies in which surgery was actually selected by fewer patients than this study would indicate?
Response. Thank you for this comment. We have edited the sentence as follows: “Previous clinical studies showed that the choice rate for POP surgery versus conservative management was slightly lower than in our study (from 33% to 44%) [17,18], while a survey-based study reported a choice rate of 48% [19].”

Comment 7. Please elaborate on the claim that the clinical scenario was "not easy to understand" (line 282). Why would this be? Was the wording poor or the language too advanced?
Response. We have better explained this sentence by editing it as follows: “we did not quantitatively investigate the readability of the text [33], but only asked qualitatively whether the clinical scenario was easily understandable, so some women may have responded without fully grasping the content

of the study.”. We meant that we did not use tools to quantify the readability of the text, as the ones described in the citation that we have added (number 33).

Comment 8. In the authors' conclusions in both the Abstract and the paper's Discussion section, the authors assert that the study results reveal a relationship between "unwarranted variation" and determinants of patient choice around prolapse treatment. However, if "unwarranted variation" involves factors NOT related to patients, then I don't think this is an accurate statement. This study seems to identify factors that influence patient preference around surgical versus non-surgical management and does NOT shed any light on sources of unwarranted variation, such as surgeon skill / preference, hospital resources, and regional medical infrastructure. Please clarify this or resolve this inconsistency.
Response. According to Wennberg taxonomy, variation is justified when it is due to the users' preferences, needs and choice. Variation is defined as unwarranted when depends on other factors than patients' choices. If a woman who receives no input would have surgery in 40% of cases, but when receiving specific inputs she would have surgery in 80% of cases, this difference could explain some of the observed variation. Clearly, these determinants are not intrinsic to the woman, but are "external" factors or inputs (travel time, advice from physicians, surgeon skills, outcomes of surgery, etc...) that are, however, investigated at the level of the individual. We have edited the Conclusion section as follows: “If a woman who receives no input would have surgery in 40% of cases, but when receiving specific inputs she would have surgery in 80% of cases, this difference could explain some of the observed variation. The determinants of such a variation have been explored at the user level, and they emerged to be related to physicians’ advice, surgical outcomes, and logistic factors such as travel distances”.

Reviewer: 3
Dr. Anne Brabers, NIVEL Netherlands Institute for Health Services Research
Dear authors,
Thank you for giving me the opportunity to review this article about practice variation, a topic in which I am highly interested. Hereafter, I will provide my feedback on the manuscript.

Comment 1. One of my main concerns is that you performed your study among women of the general population, most of them not suffering from POP. I am in doubt whether research among this group gives a good insight in patient related factors influencing the preference for surgery or not. Another main concern is whether you really get insight in individual determinants. When I think about individual determinants, I think about for example attitude, beliefs, preferences, involvement, information (see e.g. the Theory of Reasoned Action from Fishbein & Ajzen (2010)). But also factors like health literacy, and social norms (e.g. whether women in the social environment of the woman have had a surgery for POP). The last main concern is the link between the outcomes of the study and the conclusion that the variation is strictly related to physician advice. I doubt whether information on clincial outcomes depends fully on the phsycian, is there no general information available about clinical outcomes available (e.g. in a decision aid), or on the internet? Based on the fact that advice from a gynaecologist increases the choice for surgery, the authors conclude that variation is mainly due to the physician. Althoug, I understand and support this (based on other literature), I think ,this conclusion could be better substantiated. For example, the results also show that 80% opt for surgery when there is a hospital nearby, and that is not a factor related to the physicians attitude.
Furthermore, I expected that the analyses were done by region. If preferences of women do not differ between regions, the differences in regions as shown in Figure 1 (but also in whole Italy) are not related to the preferences of the women, but to other factors. Is it possible to do the analyses per region?
Response. As of the choice of enrolling women from the general population, we have responded to a similar comment, namely comment 1 by Reviewer 1. We wrote: “We decided to include women from

the general population on purpose. In fact, we believe that this allowed us to identify the determinants of choice at an early stage, before the condition occurred and the women had visited a health professional. By that time, the woman may have already received input that in some way may influence her choice of treatment. Since our main purpose was to investigate the effect of individual determinants of choice on practice variation, we inevitably had to choose this type of population to answer the research question." We have tried to better explain the rationale behind our choice in the section "Limitations and strengths."
Regarding the second concern that the Reviewer expressed, we apologize because we may have not explained properly. We did not want to investigate individual characteristics that link to an increased likelihood of choosing a surgical approach (such as those the reviewer rightly mentions), but rather wanted to investigate at the individual level the "external" factors that may influence women's choice. The term "individual" refers to the level of analysis, not the nature of the determinant. This ties in closely with the third issue raised by the reviewer, namely the link between our findings and unwarranted variation. Variation is acceptable if it depends on individual factors of choice (such as attitude, beliefs, preferences, needs, involvement, information, health literacy), but should instead be reduced when it depends on external, system-related, or logistical inputs. We have changed "individual determinants" to "individual-level determinants" throughout the manuscript to make it clearer. Is there anything else that we can do to better specify this aspect?
Finally, it is unfortunately not possible to conduct an analysis by region as we do not know the place of residence of the respondents. We only know the place from which the women responded to the online survey, but this does not necessarily coincide with residence. Unfortunately, a specific question on place of residence was not included in the survey.

Comment 2. In the abstract the authors write: 'and partly capture the individual determinants of unwarranted variation'. In the discussion the authors state that it is strictly the physician that causes variation. I should bring this more in line.
Response. We apologize if we have been unclear, but we do not seem to have stated that it is strictly the health professional who drives the variation. In the discussion, we addressed all our findings, so both the effect of the health professional's advice and the other factors that emerged to be statistically significant (outcomes of surgery, travel time, etc.). How can we improve the discussion?

Comment 3. In the introduction, line 74, the authors talk about the benefits of conservative treatment. But they are not explained, what are the benefits?
Response. These are explained in the Supplementary Table S1, which shows the questionnaire that we used. Should we also add this detail in the main text, or is it okay to leave it in the supplementary material for summary purposes?

Comment 4. On line 82, the authors already argue that practice patterns mainly depend on the surgical training. If this is the case, and I think this is among others the case, what is then the added value of looking into preferences of women to explain variation. I think the authors can argue a bit more in the introduction why looking into the role of women is interested to explain practice variation.
Response. In that sentence, we cited a study that showed, through a survey of member surgeons of the International Urogynecological Association, that much of the variation patterns in the evaluation and treatment of POP depended on academic affiliation and geographic location. We made a mistake in generalizing the results of that study. We therefore rephrased it as follows: “In fact, a survey-base study on members of the International Urogynecological Association showed that much of the practice patterns in the treatment of POP depended on academic affiliation and geographic location”

Comment 5. Regarding the methods, an important point is the selection of respondents. Why did the authors not invite only women who have POP? It appears that they even have tried to reach the target group.
Response. We think we have answered to this question in our response to your Comment 1. At this

point, including also women with a higher risk of / suffering from POP would mean reconducting the study de novo with an additional study population. However, as we explained, we chose to include women from the general population on purpose, believing that this was the right way to answer our research question.
Although we do not believe that enrolling an additional study population (women suffering from or at higher risk risk of POP) as a comparison group would help us answer the research question, we are willing to do so if the reviewers/editors deem it necessary.

Comment 6. Line 129 in the methods, wat is grade 3 of POP
Response. This was specified in the supplementary material, but we have added a short sentence to better specify what grade 3 POP is: “The questionnaire presented women with a clinical scenario of grade 3 POP (with uterus protruding out of the vagina).”

Comment 7. The authors included advantages and disadvantages in the questionnaire. On what kind of information was this based? I think this is important as the way how information is framed can also influence the choice.
Response. As explained in the method section, all information were obtained through an extensive literature review, and then validated through focus groups with health professionals. We have better specified this point in the Methods: “The survey content (e.g., specification of the clinical scenario, factors or determinants of choice, description of the two treatment alternatives) was developed through an extensive literature review… To consolidate our choices, we further carried out a focus group with a team of gynaecologists…”

Comment 8. Why is a scale from 1-99% chosen, and not 0-100? And on line 166 it is stated that it is converted in a 0-1 scale, how was this done?
Response. We chose to use a scale of 1 to 99% since a key assumption of the beta regression model is that the variable has a value between 0.01 and 0.99, and that it is never equal to 0 or 1. When we wrote "converted in a 0-1 scale," we meant the response rate was divided by 100, thus expressed in decimals. We wrote "0-1 scale" as a simplification: we have thus changed this expression to “0.01-0.99 scale”.

Comment 9. I do not understand whether everyone only saw one of the 11 questions, or saw all questions. On line 132 it is stated that each women received only one question out of 11. While on line 142 it is stated that women further received the other 10 questions in a nonrandomized manner. In table 2 it appears that only 1 question is filled out by each women. If that is the case, I think 17-23 women per factor is not that much.
Response. Each woman received only one question as the first randomized factor. After answering, on the next screen the other 10 questions were asked in a nonrandomized manner. The main outcome was the answer to the first randomized question, because the answer to the next 10 questions was susceptible to the bias due to the answer given to the first question. However, this also allowed us to perform a within-subject (in addition to between-subject) analysis, which confirmed the results of the main analysis.

Comment 10. Regarding the chosen factors, I understand that less information was available about the topic, but why did the authors search for individual factors that are related to patient choices, and adapt that to the situation? Like attitudes. For now it appear a bit a random choice of factors, and what is for example the difference between advice from a GP and advice from two physicians. Furthermore, the five gynaecologists that are consulted where all from one university, they probably think and do the same as physicians adapt to each other. Why were not physicians invited from different hospitals/regions?
Response. As we explained, factors were chosen through extensive literature review, focus groups with professionals, and public involvement (pilot survey). We tried to make the best effort with the

limited data available in the literature based on evidence from previous studies on different topics, expert opinion, and women's views. It is true that the health professionals involved were all from the same institution, but in the Tuscany region the largest pelvic floor surgery center is located at this institution. We also tried to involve professionals from other regions, but without success. Finally, we wanted to emphasize the difference between the advice of the gynecologist and the GP precisely because there were no previous data in the literature on this topic, but in Italy the figure of the GP plays a fundamental role because he or she is the health care professional in closest contact with the patient and whom the patient often trusts the most.

Comment 11. The questionnaire was piloted, in line 161-162 it is stated whether women in the pilot had other factors or determinants that were not considered by the researchers. Could the authors give some insight in whether they adapt something in the questionnare based on the pilot.
Response. Actually, the pilot phase served mainly two purposes: 1) to check the clarity of the text, and 2) to modify the text according to the women's suggestions. These suggestions were more about the description of the clinical scenario and treatment options rather than the choice of factors. In fact, the women confirmed almost all the suggested factors chosen through literature review and expert focus group. The only factor we added following women's suggestions was "advice from two physicians."

Comment 12. In the discussion, line 232/233 the authors report that other studies have lower choice rates for POP surgery. Could they explain these differences with their own study?
Response. We have addressed this issue by responding to comment 6 by Reviewer n. 2. We have edited the sentence as follows: "Previous clinical studies showed that the choice rate for POP surgery versus conservative management was slightly lower than in our study (from 33% to 44%) [17,18], while a survey-based study reported a choice rate of 48% [19]."

Comment 13. In line 234 the authors use the term 'received no input'. I think that is not the right phrase, as also those women receive input, namely from the scenario. So I think, this has to be formulated in another way.
Response. Thank you for the suggestion. We have edited the sentences: "compared with the reference question in which women were given no input no further indication than those described in the scenario" (page 10), and "when receiving no further input than those given in the clinical scenario" (page 12).

Comment 14. On line 303/305 the authors argue that guidelines can help to reduce practice variation. From the literature, we know that guidelines not reduce variation. See for example: Jong, J.D. de, Groenewegen, P.P., Spreeuwenberg, P., Schellevis, F., Westert, G.P. Do guidelines create uniformity in medical practice? Social Science & Medicine: 2010, 70(2), 209-216. Could the authors explain why they think that guidelines will work in this situation?
Response. It is true that the work of Judith D. de Jong et al. is significant, but it is a single contribution published in the Dutch context on variability in drug prescribing, and not in elective surgery. The same authors, while questioning whether the introduction of guidelines creates uniformity, conclude "the increase (in variation) was less in the cases of diagnoses for which guidelines were introduced. Guidelines, primarily, had an effect on variations in single-handed practices. The overall conclusion is that the introduction of guidelines, although it probably tempered the increase in variation, did not reduce variation." In the Introduction, the authors state "The use of clinical guidelines that give recommendations about appropriate health care is a way of reducing variation and maintaining, or improving, the quality of health care". So, the authors themselves cite various articles suggesting that guidelines could reduce variation. Finally, the difference between the Italian and Dutch National Healthcare Systems may not make the results of the study by Judith D. de Jong et al. internationally generalizable.

Comment 15. Furthermore, in the literature it is suggested that SDM (with the help of decision aids) can reduce practice variation. E.g. O'Connor AM, Llewellyn-Thomas HA, Flood AB. Modifying unwarranted variations in health care: shared decision making using patient decision aids. Health Affairs 2004; Var63-72 / Greer AL, Goodwin JS, Freeman JL, Wu ZH. Bringing the patient back in. Guidelines, practice variations, and the social context of medical practice. International Journal of Technology Assessment in Health Care 2002; 18(4):747-61 / O'Connor AM, Wennberg JE, Legare F, Llewellyn-Thomas HA, Moulton BW, Sepucha KR, et al. Toward the 'tipping point': decision aids and informed patient choice. Health Affairs 2007; 26(3):716-25 / Stiggelbout AM, Van der Weijden T, De Wit MP, Frosch D, Legare F, Montori VM, et al. Shared decision making: really putting patients at the centre of healthcare. The BMJ 2012; 344:e256. An empirical study also confirms this: Brabers AEM, van Dijk L, Groenewegen PP, et alDoes a strategy to promote shared decision-making reduce medical practice variation in the choice of either single or double embryo transfer after in vitro fertilisation? A secondary analysis of a randomised controlled trialBMJ Open 2016;6:e010894. doi: 10.1136/bmjopen-2015-010894. I think the authors have to include some literature about this topic. They mention in line 310 a well informed decision-making process, but I think they have to refer to the literature about SDM, and say something about that.
Response. Thank you very much for this contribution. We have added a paragraph in the Discussion (page 16) citing all the articles proposed by the reviewer and summarizing their contents.

VERSION 2 – REVIEW

REVIEWER	Zhuo, Yueran Mississippi State University, Marketing
REVIEW RETURNED	30-Apr-2024

GENERAL COMMENTS	Firstly, I would like to acknowledge the efforts made by the authors in addressing the comments from the first round of review. The inclusion of a control question regarding epidural analgesia during childbirth to enhance the survey's reliability, and the attempt to clarify the study rationale in the "Limitations and strengths" section, demonstrate a commitment to refining the study's methodology and presentation. These changes have undoubtedly strengthened the manuscript. 1. While appreciating the rationale provided for focusing on a general population, the study would greatly benefit from a comparative analysis with actual POP patients. Such a comparison could provide insights into the preferences of those directly affected by POP and evaluate how representative the survey results are of this specific patient group. To this end, I recommend the paper to include a group of women currently managing POP, either through surgery or conservative treatments, could enrich the findings. This additional data layer would enhance the relevance of the research to clinical practice by aligning the surveyed preferences with those of real-world patients. 2. Related to the above point, the study could further benefit from a deeper exploration of the psychological and social dynamics that shape these decisions. I suggest integrating a qualitative study component, such as interviews or focus groups, which could provide valuable insights into the reasons behind the influence of factors like physician advice or hospital proximity on patient choices. This would add depth to the quantitative data, offering a

	comprehensive understanding of how social support and information dissemination impact patient decisions in the context of POP.

REVIEWER	Collins, Sarah A The University of Chicago Hospital, Obstetrics & Gynecology
REVIEW RETURNED	02-Apr-2024

GENERAL COMMENTS	1) Lines 233-237 still do not make sense. Why differentiate factors whose ORs differ by only 0.1? Are you saying some are >2.9, some between 2.8-2.9, and some < 2.8? If so, you will need to change your symbol (<,>) usage. 2) In the new paragraph beginning 355, you discuss SDM. You should probably reference some articles that have specifically addressed SDM in women choosing surgery for POP, such as Female Pelvic Medicine & Reconstructive Surgery 27(2):p e309-e314, February 2021 AND Female Pelvic Med Reconstr Surg . 2015 Jul-Aug;21(4):231-5 3) What is meant by "specific scientific inputs" (line 386)? This might help me to understand the sentence added here.

REVIEWER	Brabers, Anne NIVEL Netherlands Institute for Health Services Research
REVIEW RETURNED	26-Apr-2024

GENERAL COMMENTS	Dear authors, Thanks for giving a reaction to the comments raised during my first round of review. I think the article has improved, however, I am still in doubt of the study population is the most suitable population. As of now, you included some extra information about the rationale for the study population in the limitations. I think it will help if you also include some rationale for this decision in the methods. Furthermore, the response rate is now added, and is with 89% really high. I am in doubt of the authors can say that 250 women are invited. As they describe that the link was spread using social media and that women might forward the link. I think that with this open link strategy, it is difficult to determine an exact respons percentage. Although the explanation about the individual level determinants helps me in understanding what the authors mean by individual determinants, I still think it is a confusing term. Also are not all factors at the individual level, as the availability of a hospital nearby or not is more on the organisational/meso level. It are 'determinants influencing the choice of the individual women'. And personally, I would use that phrasing to overcome any confusion about what is mentioned with individual. The authors have no information about the region of the woman. In reaction to one of the other reviewers they mention that they got back to the respondents to ask a control question. Why decided the authors to only ask a control question, and not also to the region of the women? And when going back, did all 222 women respond to the control question? Regarding previous comment 3, for me it would be helpful to have one or two sentences added with the benefits.

	Regarding previous comment 9, I understand the explanation of the authors, but I think in the manuscript it is still confusing. On line 142 it is written that: Each woman randomly received only one question out of 11. And on lines 150-153 it was written as: The first question was always randomized to avoid response bias, related to the inevitable influence by the answer to the first question on the answer to the following questions. However, women further received the other 10 questions in a nonrandomized manner. I think the authors could write this more in line with each other to avoid confusion. Regarding previouw comment 11: I think it is good to add in the methods that one factor was added by women in the pilot study. And than also indicate which factor was added (advice from two physicians)

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Dr. Sarah A Collins, The University of Chicago Hospital
Comment 1. Lines 233-237 still do not make sense. Why differentiate factors whose ORs differ by only 0.1? Are you saying some are >2.9, some between 2.8-2.9, and some < 2.8? If so, you will need to change your symbol (<,>) usage.
Response. You are right. Thank you for the suggestion. We have reworded the paragraph by removing any reference to OR values (page 11).

Comment 2. In the new paragraph beginning 355, you discuss SDM. You should probably reference some articles that have specifically addressed SDM in women choosing surgery for POP, such as Female Pelvic Medicine & Reconstructive Surgery 27(2):p e309-e314, February 2021 AND Female Pelvic Med Reconstr Surg. 2015 Jul-Aug;21(4):231-5
Response. Done (page 17, line 369). Thank you.

Comment 3. What is meant by "specific scientific inputs" (line 386)? This might help me to understand the sentence added here.
Response. We meant the determinants of choice analyzed in our article (physicians' advice, surgical outcomes, travel times, …). We have modified the text accordingly (page 18).

Reviewer: 3
Dr. Anne Brabers, NIVEL Netherlands Institute for Health Services Research
Comment 1. Dear authors, Thanks for giving a reaction to the comments raised during my first round of review. I think the article has improved, however, I am still in doubt of the study population is the most suitable population. As of now, you included some extra information about the rationale for the study population in the limitations. I think it will help if you also include some rationale for this decision in the methods.
Response. Done (please, see page 6). Thank you for your suggestion.

Comment 2. Furthermore, the response rate is now added, and is with 89% really high. I am in doubt of the authors can say that 250 women are invited. As they describe that the link was spread using social media and that women might forward the link. I think that with this open link strategy, it is difficult to determine an exact response percentage.

Response. Actually, we know that 250 were contacted because the Software we used to administer the questionnaire (Qualtrics) shows how many people opened the link (250 precisely). Of these, however, only the 222 women included in the study completed the questionnaire. Obviously, we have no information on how many women received the link without even opening it. We have modified the first sentence of the “Results” section accordingly (page 10).

Comment 3. Although the explanation about the individual level determinants helps me in understanding what the authors mean by individual determinants, I still think it is a confusing term. Also are not all factors at the individual level, as the availability of a hospital nearby or not is more on the organisational/meso level. It are 'determinants influencing the choice of the individual women'. And personally, I would use that phrasing to overcome any confusion about what is mentioned with individual.
Response. Thank you for the suggestion. We have modified the text by using the phrasing recommended by the reviewer throughout the manuscript.

Comment 4. The authors have no information about the region of the woman. In reaction to one of the other reviewers they mention that they got back to the respondents to ask a control question. Why decided the authors to only ask a control question, and not also to the region of the women? And when going back, did all 222 women respond to the control question?
Response. Sorry, we did not specify that just 203 women answered the control question. We have modified the text accordingly (page 11). Also, we re-contacted the women through Qualtrics and asked a specific question about the region of residence. We have modified the text accordingly (Result section, page 10) and included a specific Supplemental Table S3 in the Supplementary Material.

Comment 5. Regarding previous comment 3, for me it would be helpful to have one or two sentences added with the benefits.
Response. Done. Please, see page 8, lines 140-142.

Comment 6. Regarding previous comment 9, I understand the explanation of the authors, but I think in the manuscript it is still confusing. On line 142 it is written that: Each woman randomly received only one question out of 11. And on lines 150-153 it was written as: The first question was always randomized to avoid response bias, related to the inevitable influence by the answer to the first question on the answer to the following questions. However, women further received the other 10 questions in a nonrandomized manner. I think the authors could write this more in line with each other to avoid confusion.
Response. Thank you for noticing this inconsistence. We have modified lines 144-145 (page 8) to align them with the following sentence at lines 153-155.

Comment 7. Regarding previouw comment 11: I think it is good to add in the methods that one factor was added by women in the pilot study. And than also indicate which factor was added (advice from two physicians)
Response. Done. Thank you for the suggestion. See page 9, line 177.

Reviewer: 1
Ms. Yueran Zhuo, Mississippi State University
Firstly, I would like to acknowledge the efforts made by the authors in addressing the comments from the first round of review. The inclusion of a control question regarding epidural analgesia during childbirth to enhance the survey's reliability, and the attempt to clarify the study rationale in the "Limitations and strengths" section, demonstrate a commitment to refining the study's methodology and presentation. These changes have undoubtedly strengthened the manuscript.

Comment 1. While appreciating the rationale provided for focusing on a general population, the study would greatly benefit from a comparative analysis with actual POP patients. Such a comparison could provide insights into the preferences of those directly affected by POP and evaluate how representative the survey results are of this specific patient group.
To this end, I recommend the paper to include a group of women currently managing POP, either through surgery or conservative treatments, could enrich the findings. This additional data layer would enhance the relevance of the research to clinical practice by aligning the surveyed preferences with those of real-world patients.
Response. We want to thank the reviewer for this suggestion. However, we feel that addressing this request would involve a considerable effort of time and resources without much contribution to answering our research question. While we acknowledge that such a comparative analysis could provide interesting insights, we would like to point out that the key objective of the study is to demonstrate a potential link between individual women's decision making and the presence of unwarranted variation in POP surgical practice. This article would represent the last chapter of the first author's PhD thesis, which is focused on variation in POP surgery. Administering the questionnaire to women who suffer from POP would mean enrolling them at the time they visit a health professional for these issues. By that time, women may have already received the opinion of a health professional, hence somehow biasing their answers to our questionnaire. Also, by choosing a hospital or outpatient urogynecological clinic they may have already made a choice related to, for example, travel distances, or reputation, or advice from doctors, friends, relatives. We think that the reviewer's suggestion, i.e. to also administer the questionnaire to a population of women with POP, could be accomplished in a following study with a different objective. A future study could be designed differently using the results of the current analysis. For example, a Discrete Choice Experiment could be conducted, rather than using the same questionnaire already adopted (which in any case should be modified to make it consistent with a different study population). In conclusion, we believe that this additional analysis suggested by the reviewer could be avoided.

Comment 2. Related to the above point, the study could further benefit from a deeper exploration of the psychological and social dynamics that shape these decisions. I suggest integrating a qualitative study component, such as interviews or focus groups, which could provide valuable insights into the reasons behind the influence of factors like physician advice or hospital proximity on patient choices. This would add depth to the quantitative data, offering a comprehensive understanding of how social support and information dissemination impact patient decisions in the context of POP.
Response. While we appreciate this feedback, we believe that this kind of in-depth study would not be feasible. Indeed, we would not have the opportunity to individually contact some of the women who took part in the study and ask them to participate in an interview or focus group due to privacy issues. This would require approval from an ethics committee, inevitably lengthening the time to complete this additional analysis. Alternatively, if we selected other women than those who participated in the survey for an interview or focus group, it would be impossible to link the survey results with the qualitative questions and thus investigate the psychological and social dynamics that influence women's decisions. In general, we believe that this request would be too complex and burdensome for the type of study we have designed from the beginning and could therefore be satisfied in a subsequent study.

VERSION 3 – REVIEW

REVIEWER	Collins, Sarah A The University of Chicago Hospital, Obstetrics & Gynecology
REVIEW RETURNED	29-May-2024

GENERAL COMMENTS	Well done! I hope to see your article in print soon.

VERSION 3 – AUTHOR RESPONSE

Reviewer: 2
Dr. Sarah A Collins, The University of Chicago Hospital
Comment 1. Well done! I hope to see your article in print soon.
Response. Thank you very much for your valuable comments that allowed us to improve our article.