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)	Development of a prediction model to aid primary care physicians in early identification of women at high risk of developing endometriosis: a cross-sectional study
AUTHORS	Verket, Nina; Sørum Falk, Ragnhild; Qvigstad, Erik; Tanbo, Tom; Sandvik, Leiv

VERSION 1 – REVIEW

REVIEWER	Sanjay K. Agarwal, MD UC San Diego USA
REVIEW RETURNED	01-Apr-2019

GENERAL COMMENTS	The manuscript addresses an important problem in endometriosis care and that is delay in diagnosis, part of which is due to suboptimal disease awareness. The authors construct a model to identify women with endometriosis based on 6 test predictive questions. They also assess other basic symptom characteristics including presence of dysmenorrhea, pelvic pain, dysuria, dyschezia, fatigue nausea, irregular bleeding and irregular bowel movements in the four weeks prior to questionnaire administration. The Aspects of the manuscript needing clarification are: 1. On what basis were these 6 study questions (candidate predictors) selected for the model? Much more detail is needed here. 2. Apart from age at menarche, the other 5 candidate predictors were significantly more common in women with endometriosis. The strongest of these were frequent absenteeism from school due to painful menstruations and positive family history of endometriosis. These rendered an AUC of 0.83%. It is not clear whether the candidate predictors add any further diagnostic ability to the basic symptom data (dysmenorrhea, pelvic pain, dysuria, dyschezia, fatigue nausea, irregular bleeding and irregular bowel movements), which also was very different between women with endometriosis and those in the general community (Table 1). Please clarify. 3. For a diagnostic test of endometriosis to be needed and clinically useful, the control group should have pain symptoms as the endometriosis group. In this study, the control group had significantly less pain than those in the endometriosis group with a P<0.001 for each of dysmenorrhea, pelvic pain, dysuria, dyschezia, fatigue nausea, irregular bleeding and irregular bowel movements). There

	may not be much of a diagnostic dilemma if the patent does not have endometriosis symptoms. 4. What led to those in the endometriosis group having a laparoscopy and hence being diagnosed with endometriosis? If they presented, for example, with symptoms of endometriosis such as pelvic pain, frequent absenteeism from school due to painful menstruations and positive family history of endometriosis and these symptoms led to a laparoscopy and diagnosis of endometriosis, then it is no surprise that the endometriosis group has a history of pelvic pain, more frequent absenteeism from school due to painful menstruations and positive family history of endometriosis. An ideal control group would have had a laparoscopy and found not to have endometriosis.

REVIEWER	Fauconnier A. EA 7285 Research Unit 'Risk and Safety in Clinical Medicine for Women and Perinatal Health', Versailles-Saint-Quentin University (UVSQ), 78180 Montigny-le-Bretonneux, France Department of Gynecology and Obstetrics, Centre Hospitalier Intercommunal de Poissy-Saint-Germain-en-Laye, 78300 Poissy, France
REVIEW RETURNED	29-May-2019

GENERAL COMMENTS	This is a cross-sectional study aiming to develop and validate a prediction model for endometriosis that could be use for primary care. I would like to congratulate the authors because the development of prediction model for endometriosis is a priority of the research in the field. However, I have several concerns. 1. Could the author provide a Flow chart to identify the number of patients who respond to the questionnaire? Moreover, it would be important to well identify the number of missing data (percentage) and the reason of missing data. Indeed, in this kind of study it is crucial to state the exact number of missing data according to each response of the questionnaire. 2. Six candidate predictors were used in this study. Could the author specify how they choose these candidates? Is it a choice of a scientific committee? Please give more details. 3. The statistical analysis is well perform and in line with the TRIPOD guidelines. 4. I have no experience of lasso regression analysis. Therefore it could be interesting to have a statistical reviewer. 5. Did the author could provide more details on the definitions of the 5 five categories for the candidate predictors 2 and 4. For example, how they define rarely for the use of painkiller (for example: less than 3 times per week). 6. Missing data: It is a crucial point to improve this paper. The authors should give much more details on missing data. How they manage these missing data. Moreover, it could be important to perform a sensitivity analysis with a complete case analysis. For example, page 8 line 44, for the "family history", 15 did not answer at all. 7. Could the authors justify why they choose to consider the variables "severe dysmenorrhea" and "absenteeism from school" as continuous. In table 1, these two variables seem to be categorical ("never", "rarely"…). Is this choice modify the results of the study. 8. In my opinion, the inclusion of patients from endometriosis association could overestimate the proportion of positive response

	for family history. Could you comment this point? 9. Concerning ERI_2. Why the authors choose the cut off of 12, 19, 26 and 33? 10. Page 5 line 5 and 7: is 1500 women or 1050 women?

VERSION 1 – AUTHOR RESPONSE

- **Response to comment 1 of Reviewer 1:**

We agree. The rationale behind the choice of the six candidate predictors has been included in the paragraph “Candidate predictors” under the section of “Participants and methods” in the revised manuscript.

- **Response to comment 2 of Reviewer 1:**

We agree that presentation of data in table 1 is not optimal. Table 1 has therefore been split into two tables: Table 1 now describes recent characteristics (experienced at any time during the 4 weeks prior to answering the questionnaire) of the participants. The new table 2 describes adolescent characteristics and family history of endometriosis of the participants.

The aim of the study was to develop a prediction model to aid primary care physicians in early identification of women at risk of developing endometriosis. We therefore chose, as candidate predictors, variables that would have been availableto the respondents already during adolescence – keeping in mind that the mean age of the respondents was above 30 years. The basic characteristics data describing recent symptoms (dysmenorrhea, pelvic pain, dysuria, dyschezia, fatigue, nausea, irregular bleeding, and irregular bowel movements) were not included as candidate predictors because recent symptoms would not be suitable variables to include in a prediction model for earlier identification of women at risk of endometriosis from a much younger age.

- **Response to comment 3 of Reviewer 1:**

We agree that the prediction models presented in our study are not suitable for use in gynecological surgical departments to improve selection of women with symptoms suggestive of endometriosis for diagnostic laparoscopic surgery with the aim to reduce the personal and institutional costs associated with unnecessary procedures. If this were our objective, we agree that we should have used a control group consisting of women with symptoms suggestive of endometriosis who were found not to have endometriosis during laparoscopy. However, our prediction model development study differs in having as its main focus early identification of young women in the general population at high risk of developing endometriosis. We therefore recruited our control group from a randomly selected sample from the general population.

- **Response to comment 4 of Reviewer 1:**

A key point with respect to our study is that, although symptoms suggestive of endometriosis are more frequent among women with endometriosis, they are common enough among women in the general population to make diagnosis challenging, leading to late referrals to specialists by primary care physicians, c.f. the longer diagnostic delay in primary care

suggested by several studies. To develop a prediction model to aid primary care physicians in early identification of women from the general population at risk of developing endometriosis and thereby reduce diagnostic delay, the control groupspan style="font-family:Arial; color:#222222"> in our view had to be from the general population.

- **Response to comment 1 of Reviewer 2:**

Thank you for your comment. A supplementary flow chart has been added to the revised manuscript. Table 1 has been split into two tables (table 1 and table 2). The exact number of missing data (and the percentage) for each candidate predictor has been added to the new table 2.

The reason for missing data has been added to the paragraph “Candidate predictors” under the section of “Results” in the revised manuscript. Blank responses were described as missing. In the control group, six participants skipped an entire page of the questionnaire (including the candidate predictors) most likely by error, resulting in missing responses for all the candidate predictors.

- **Response to comment 2 of Reviewer 2:**

Thank you for your comment. The authors chose the candidate predictors based on three criteria, the description of which has been added in the paragraph “Candidate predictors” under the section of “Participants and methods” in the revised manuscript. A representative of the Norwegian Endometriosis Association assessed and confirmed legibility of the questions (candidate predictors).

- **Response to comment 3 of Reviewer 2:**

We appreciate your comment. Thank you very much.

- **Response to comment 4 of Reviewer 2:**

The explanation of why we employed lasso regression analysis has been slightly expanded in the paragraph “Statistical analysis” under the section “Participants and methods” in the revised manuscript.

- **Response to comment 5 of Reviewer 2:**

Because this study, to the best of our knowledge, is the first to attempt screening tool development for use in primary care, we were reluctant to go too far in detailing the response categories. Having no data to suggest appropriate further specification of the response categories of the candidate predictors, we also considered the risk of using inappropriate specifications. Several studies report fewer response categories for the variable “absenteeism from school due to dysmenorrhea” (for example yes/no). However, in an external validation study of the prediction models, we clearly see the need for expanding the response categories (never, rarely, sometimes, often, and always) further. The response categories should be specified to offer better user instructions.

- **Response to comment 6 of Reviewer 2:**

> Questionnaires with missing data for any of the six candidate predictors, except for the candidate predictor “Family history of endometriosis”, were excluded from regression analyses (logistic and lasso regression analysis). Thus, data from 154 women with endometriosis and 145 without endometriosis were included in the original analyses. The candidate predictor “Family history of endometriosis” was included as a dichotomous variable (categorized as “Yes” versus “No”/“I don’t know”/“Irrelevant”/missing). A dichotomous categorization was chosen for this candidate predictor to design a screening tool that would handle the real-life response categories “I don’t know” and “Irrelevant” (for example if adopted). Further, for the control group, considering no one answered “I don’t know” and 15 participants gave a blank response, we thought it likely that many blank responses in reality were comparable to the response “I don’t know”. We therefore decided that it would be best to include all response categories for the candidate predictor “Family history of endometriosis” in the analyses.

However, in line with your suggestion, we decided to perform a complete case analysis, excluding all missing data for all candidate predictors, including “Family history of endometriosis”.
Only questionnaires with “Yes” or “No” responses for the candidate predictor “Family history of endometriosis” were included. Questionnaires with “I don’t know”, “Irrelevant”, and missing (blank) responses for this predictor were excluded. This gave a sample of data from 142 women with endometriosis and 130 women without endometriosis for regression analyses. Regression analyses with this sample gave almost identical results as the original regression analyses performed with data from154 women with endometriosis and 145 without endometriosis. The results have been included in the revised manuscript as a supplementary table.

VERSION 2 – REVIEW

REVIEWER	Fauconnier Arnaud CHI Poissy-St-Germain Poissy, France
REVIEW RETURNED	19-Sep-2019

GENERAL COMMENTS	The present manuscript has been improved partially by the authors and most of the remarks of the reviewers where correctly answered. I think there is valuable data in its present forms which must be therefor accepted. I have nonetheless two remarks: I don't think there is a need to present both regression models in the manuscript. There is, in general no value to perform several distinct statistical modeling models to develop clinical prediction models. It is a part of the preliminary step of the procedure to choose the appropriate model (as it is important to pre-select the appropriate predictors, see below). Using several statistical regression models in the same time, may create overestimation of the prediction performance because you will finally chose the best fitting model and not the best set of predictors. As you finally opted for the logistic regression, I would therefore remove the lasso regression results. This would make your main result more in light and also avoid long explanations that finally bring nothing interesting. My second concern is about the questionnaire you did use. How were the questions constructed, did they came from previous validated questionnaire, if not How can you take into account their comprehensibility (as you rightly stated). Please give the exact wording of the questions that were use.

VERSION 2 – AUTHOR RESPONSE

Response to Reviewer:

Thank you very much for your remarks.

Regarding the first remark, we agree with the Reviewer that it is in general unnecessary to present more than one statistical analysis for the same purpose, and that presenting multiple analyses would unnecessarily confuse the reader. However, lasso regression analysis is recommended in the TRIPOD checklist for developing prediction models. Lasso regression analysis is recommended because it counteracts overestimation of prediction model performance. At the same time, logistic regression analysis is one of the most commonly used analyses for prediction model development and familiar to many readers. The fact that the two statistical approaches (logistic and lasso regression analysis) reveal similar results in our study, which was not an obvious outcome, may be viewed as strengthening our findings. Moreover, because our prediction models are best estimates based on our sample, in an external validation study it would be a strength to have two models, thereby increasing the chances of finding the model best fit for real life.

Although not clearly stated, our suggestion to opt for the simplest prediction model pertains to the future. At this stage of prediction model development, such a suggestion would be premature. We have therefore adjusted the abstract (by adding the results of the lasso regression analysis) and adjusted the discussion (by replacing the last sentence of the first paragraph “Statement of principal findings”) to remove any implication of favoring one prediction model over the other.

Regarding the second remark on the use of a validated questionnaire, we have given the exact wording of the questions that were used, translated into English by the authors, in the revised manuscript, section “Participants and methods”, paragraph “Candidate predictors”. The questions on which this study is based are not taken from a validated questionnaire, because, to the best of our knowledge, no standard questionnaire for developing prediction models for endometriosis is available. The questions were made by the authors, with a view to collect accurate information on the candidate predictors chosen. The questionnaire, including the candidate predictors, was also assessed by a representative of the Norwegian Endometriosis Association for readability prior to survey administration.

A relatively high data completeness may indicate that the participants found the questions (candidate predictors) meaningful and relatively easy to answer. However, missing data, as pointed out by the Reviewer previously, was slightly higher for the candidate predictor “family history of endometriosis” than for the other candidate predictors. Regarding the candidate predictor “family history of endometriosis”, a few participants with blank responses (missing data) had written the comment “I don’t know” in the answer field. These missing data were categorized as “I don’t know”. Although this detail has no effect on the analyses performed and the findings, this distinction has been clarified in the revised manuscript. This distinction may also explain the slightly higher number of missing data for this candidate predictor. Of eight participants with endometriosis with blank responses to “family history of endometriosis”, seven had written the comment “I don’t know” in the answer field. Although none of the participants without endometriosis had provided similar comments, it seems likely that some of the blank responses represent participants simply not knowing. As added to the paragraph “Statement of principal findings” of the discussion in the revised manuscript, this suggests that “I don’t know” should be included as a response category (in addition to “Yes”, “No”, and “Irrelevant”) for this predictor in future studies.

Our questionnaire study was anonymous. Thus, we did not have access to medical records of participants or family members of participants. We could therefore, for example, not validate presence

or absence of family history of endometriosis.

As stated in the manuscript, external validation of the prediction models is important and necessary before model implementation. The questions on which the prediction models are based, would be part of such an external validation.