

Peer Review File

Article Information: <https://dx.doi.org/10.21037/tcr-21-1919>

Review comments

Comment 1: First, the title is not clear. In fact, the authors developed and tested the predictive performance of a predicting model for BCS (failure and success). Please make the title more clear to reflect the work done by the authors.

Reply 1: Thank you very much for your comment. We changed the title into“ Preoperative MRI features associated with failed breast conserving surgery: construction of a predictive model” .

Changes in the text: Page 1, line 1-3

Comment 2: Second, English language of this paper is poor, which needs professional editing after extensive revisions.

Reply 2: Thank you for suggestion. We sent the paper to the professional editing in order to improve the English language, and provided the English editing certificate.

Comment 3:Third, in the abstract, please indicate the strengths of predictive models based on MRI in the background. In the part of results, please describe variables included in the predicting model. I suggest the authors to report sensitivity and specificity of the predicting model. The conclusion should be made with cautions due to the small sample size and retrospective design.

Reply 3: Thank you very much for your comments. We revised the abstract accordingly.

Changes in the text: Page 4.

Comment 4:Fourth, in the introduction part, please focus on the need for predicting the outcome of BCS, overview of known predictive models for BCS and their limitations, the strengths of predictive models based on MRI, and the potential clinical significance of this research topic.

Reply 4:

Changes in the text: Page 5, line 119-123, Page 6, line 141-145,153-156,159-163.

Comment 5:Fifth, the study has two basic methodology issues, first, the sample size is inadequate, particular failed BCS, and, second, there is no external independent validation sample. The authors can not use the sample used to develop the model to validate the performance of the predictive model, which often overestimate the performance. So I think the current study can not answer the question of whether the

predictive model is appropriate for predicting. This is a failed study.

Reply 5: Thank you very much for your insightful comments. Generally, 1:1 matching are used in researches; however, because there are few cases of breast conserving failure in this study, and we hope to increase the sample size as much as possible to increase the study power. Generally, the study power doesn't increase if the ratio is more than 1:4. Therefore, we initially consider the 1:4 matching; but some failed BCS cases cannot be matched to 4 samples, according to matching factors of age, neoadjuvant therapy and hormone receptor expression. Thus, we finally choose the 1:3 matching.

Comment 6: Finally, in the statistics, please consider to report sensitivity and specificity. As I said above, this study can not answer the question of performance of the predictive model. I suggest the authors to focus on factors associated with failed BCS, but the small sample still limits the analysis.

Reply 6: Thank you very much for your insightful comments. We accepted your comment and revised the purpose of the study as “this study proposed to explore preoperative MRI features associated with failed BCS and construct a MRI-based predictive model for failed BCS, in order to help clinicians make more accurate determination for BCS.”(See Page 6,line 64-166).

Meanwhile, we divided patients chronologically into training group(including 15 patents with failed BCS and 45 patients with successful BCS) which was used for model construction, and testing group(including 15 patents with failed BCS and 45 patients with successful BCS) which was used for model validation.(See Page 7,line 188-191).We also revised related statistical analysis method(See Page 9,line 253-258).

We also report sensitivity, specificity, PPV, NPV and total accuracy for training group and testing group in the result part(See Page 11,line 294-306). We revised Table 3 and Figure 3 accordingly.

In addition, the small sample size and lack of external independent validation samples are in fact main limitations of this study, we revised the limitation part(See Page 12-13,line 344-350).