

Peer Review File

Article information: <https://dx.doi.org/10.21037/tp-23-309>

Reviewer A

The manuscript describes the novel application of a machine-learning-based method using features derived from CT-scans to diagnose primary immunodeficiency (PID) in children with pulmonary tuberculosis (PTB). Although the sample-size is very small for this type of analysis, the study is well done and the results show promise as a useful diagnostic.

An important limitation regarding the usefulness of such a tool relates to the low incidence of PID in PTB. The authors do not state what the incidence is in their population and this omission should be corrected. However, it is unlikely to be anywhere near as high as in their dataset. As a result of this low prior probability of PID and the moderate performance of the current model, the posterior probability of PID, given a positive test will be very small, resulting in a much higher false positive rate than the results suggest.

This is an aspect that the authors should acknowledge and discuss, giving performance estimates based on actual prevalence.

Comment:

Reply: There were no accurate statistical data of the incidence of PIDs in PTB at present. PID is still a rare disease, and it is rare both for the exact incidence and diagnostic rate, because the diagnosis requires genetic testing. Our institution is the largest children's medical center in southwest China, the largest PID clinical diagnosis and treatment center in China, and ranked third in comprehensive strength in China. Besides, the special geographical environment and undeveloped economic level lead to a higher infection rate and incidence of pulmonary tuberculosis in children here. We believe that there are many patients of PIDs in PTB not be diagnosed in children. So, the diagnostic positive rate will be higher than the national even global average. And, such a tool will be practical and helpful in southwest China and other undeveloped areas.

Changes in the text: we added the illustration about the incidence of PIDs in PTB.(see Page 3, line 68)

Textual corrections and comments

Abstract methods

line 24. 'The cohort was randomly stratified sampling into' should read 'The cohort was randomly divided into'

Reply: we agree the advice and have modified our text.

Changes in the text: The samplings was randomly divided into training and testing groups according to a 3:1 ratio.(see Page 1, line 24-25)

Since the classes are imbalanced, it would be preferable to report the F1 score rather than the accuracy.

Reply: Yes. We have recognized the problem after reviewer comment. F1 score is one

of the indicators for evaluating the model in our study.

Changes in the text: we added the illustration (see Page 1, line 31, Page 2, line 35-38, and Page 15, line 243-244)

Introduction

line 61. The claim that 'newborns and infants who receive routine Bacillus Calmette-Guerin (BCG) vaccination ... are at a higher risk of Mtb infection' does not make sense. It would be more correct to say that children with PIDs are at very high risk of BCGosis. Putting the risk of Mtb infection and BCGosis together is confusing and the text would read better if the two were separated.

Reply: we agree the advice and have modified our text.

Changes in the text: Analogously, newborns and infants with PIDs are at a high risk of BCGosis when receive routine Bacillus Calmette-Guerin (BCG) vaccination in areas where TB is endemic.(see Page 3, line 60-63)

line 77. This should be corrected to read: 'Therefore, identification of the presence or absence of PIDs in children'

Reply: we have modified our text as advised.

Changes in the text: Therefore, identification of the presence or absence of PIDs in children with PTB is crucial for their clinical treatment strategies.(see Page 3, line 77-79)

Methods

line 109 should read: 'We retrospectively **enrolled** the children'

Reply: we agree the advice and have modified our text.

Changes in the text: We retrospectively **enrolled** the children with the tested immunity status for pulmonary tuberculosis who visited our hospital from January 2015 to December 2022, and the chest CT data before treatment of all the children were collected. (see Page 4, line 109-112)

line 120. (and other sections) It is not clear what the authors mean by 50% discount cross-validation as this does not appear to be a standard method. This method should either be explained (or referenced) or another term should be used if this is in fact an error.

Reply: The describe of 50% discount cross-validation is error writing with 5-fold cross-validation in text owing to personal mistakes. So, we have modified our text.

Changes in the text: Due to the limited sample size, we conducted model training with **5-fold** cross-validation in the training set and internal validation. (see Page 5, line 119-121)

Figure 2 describes standard cross-validation and is therefore unnecessary.

Reply: Figure 2 is a flow chart of the data grouping according to the standard of **5-fold** cross-validation. It is helpful for readers to understand the grouping at first.

Changes in the text: Delete as advised.

lines 168-170. 'The hyperparameters was confirmed by a 50% discount cross-validation, sought and determine the optimal alpha value was automatically.' This sentence should probably read: 'The hyperparameters, including the alpha value, were determined automatically by 50% discount cross-validation.'

Reply: we agree the advice and have modified our text.

Changes in the text: The hyperparameters, including the alpha value, were determined automatically by 5-fold cross-validation.(see Page 7, line 168-170)

lines 198 to 200. The definitions of the various performance metrics are standard and should be removed.

Reply: Yes. The definitions of these performance metrics are standard. It is described only to help the reader to understand its meaning. It can be removed.

Changes in the text: The performance of the classifier was evaluated using the following indicators in this study, including sensitivity, specificity, accuracy, precision, recall, and F1 score.

Some contents be removed in the table.(see Page 2, line 48 and Page 8, line 198-200)

TP	True positives
TN	True negatives
FP	False positives
FN	False negatives
R	Recall

Results

line 232 should read: 'A logistic regression model'

Reply: we have modified our text as advised.

Changes in the text: A logistic regression model was used to classify the immune function status of children with PTB, and the classification threshold of the model was set to 0.5. (see Page 10, line 232-233)

Discussion

line 343. The sensitivity in the test set is not higher as stated, it is slightly lower than in the training set.

Reply: Maybe we describe inaccurately in text. Our means was that the value of sensitivity in the test set is higher than the value of specificity.

Changes in the text: The performance of the current model was moderate in the test set which the sensitivity was 0.722 and the specificity was 0.692. Besides, the value of F1-score in the test set was expected which indicated a high accuracy of the model. (see Page 15, line 243-244)

Reviewer B

Figures and tables

(1) The time span in the main text is inconsistent with Figure 1. Please check.

137 tuberculosis who visited our hospital from January 2015 to December 2022, and the
138 chest CT data before treatment of all the children were collected. The patient

(2) Please provide all figures (except Figure 1) in standalone tiff/png/jpg files with higher resolutions and with correct figure name.

(3) Please check through all figures and tables and make sure that **all abbreviations have been defined in each legend**. For example, please provide the full names of “CT” in Figure 1 and “ROI” in Figure 2.

You may consider using this format:

E.g. Figure 2. XXXXX. (A) xxxx; (B) xxxx; (C)xxxx. **Abbreviations:** xxx, xxx; xxx, xxx; ...

(4) The citation of Figure 2 is missing in the text, please check and revise.

(5) There is no Figure 7 in the manuscript, please check and revise.

(6) Figure 7 is too vague. Please resubmit in higher resolution in jpg/tiff format to us.

(7) The legend seems to be inconsistent with the text of Figure 3, please **check through** and revise.

- **Figure 3**'s legend:

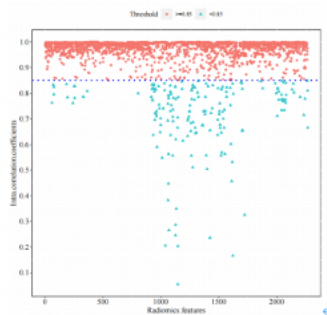


Fig.3 The ICCs distribution of the radiomics features extracted.

- The **ICCs** distribution of the 2,264 features is shown in **Figure 4**.

(8) Please provide Table 2 in editable WORD format, not figure.

Reply:

(1) Because we added some enrolled data of cases during the study from May 2022 to Dec 2022 but neglected the modification in Figure 1.

Changes in the text: We have modified in Figure 1.

(2) Yes.

Changes in the text: We have provide all figures in standalone PNG files with higher resolutions and with correct figure name attached in the replied e-mail.

(3) Yes.

Changes in the text: We have added the abbreviations defined in each legend of figures and tables.

(4) Yes.

Changes in the text: We have changed in the text.(see Page 6, line 148)

(5) and (7) Yes. Figure 7 in the text was Figure 6 actually.

Changes in the text: We have changed in the text (see Page 10, line 274) and resubmitted

a Figure with higher resolution in PNG format attached in the replied e-mail.

(8) Yes.

Changes in the text: We have changed in the text.(see Page 8, line 219)

(9) Yes.

Changes in the text: We have provide Table 2 in editable WORD format attached in the replied e-mail.