PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	A systematic review of prediction models for pulmonary tuberculosis treatment outcomes in adults
AUTHORS	Peetluk, Lauren; Ridolfi, Felipe; Rebeiro, Peter; Liu, Dandan; Rolla, Valeria; Sterling, Timothy

VERSION 1 – REVIEW

REVIEWER	Benjamin Momo Kadia Liverpool School of Tropical Medicine, UK
REVIEW RETURNED	28-Nov-2020

GENERAL COMMENTS	This is a well conducted, detailed and important review that serves to inform providers of TB treatment and care. Please, consider the following comments to improve on the manuscript. 1) Title: Please, make the title more concise by specifying 'drug-sensitive pulmonary tuberculosis in adults'. 2) Abstract: Please, also consider including the above specification in the abstract. The methods should also provide the method of data synthesis. 3) Introduction: The authors assert that ‘to date, there has not been a formal synthesis or quality assessment of existing prediction models for TB treatment outcomes, which is essential to determine which models should inform clinical practice’. This statement is untrue. Van Wyk and colleagues reported findings of a systematic review of prediction models for prevalent pulmonary tuberculosis in 2017. It is strange that the authors did not pick up this study in their search. Because of this deficit, the authors have been unable to perform a critical appraisal of this important preliminary study in the introduction or a comparative analysis with their study in the discussion. 4) Please, check the attached CHARMS checklist to ensure that it is completed.

REVIEWER	Azeez Adeboye University of Fort Hare South Africa
REVIEW RETURNED	29-Nov-2020

GENERAL COMMENTS	1. Include your PROSPERO identification number for this systematic review. 2. Is there any number of people who review the quality assessment?

	3. What are the review implications of your study?

REVIEWER	Dr. Toyin Togun, MD PhD London School of Hygiene & Tropical Medicine (LSHTM), London, United Kingdom, and Medical Research Council Unit The Gambia at the London School of Hygiene & Tropical Medicine (MRCG at LSHTM), Fajara, The Gambia.
REVIEW RETURNED	05-Dec-2020

GENERAL COMMENTS	Thank you for inviting me to review this article that aimed to systematically review and critically evaluate 37 prediction models of TB treatment outcomes from 33 studies. The authors carried out a robust search and used modern guidelines in evaluating the selected studies. Importantly, they crucially assessed the quality of the studies and risk of bias across population and analysis domain. While the authors highlighted some common predictors of TB treatment outcomes, they emphasised the limitations in the clinical use of the prediction models of TB treatment outcome based on their evaluation. The paper is well written with clear rationale and unambiguous hypothesis. They also made appropriate inferences in the discussion from the results of their analysis and evaluation. However, I have a couple of minor comments that I wish the authors will provide some clarifications: In lines 48 - 51 (page 7), the authors stated that "articles that included drug-susceptible and drug resistant-cases or a combination of adult and children were included." The factors that influence the TB treatment outcomes, and thus prediction models, in subjects with drug-resistant TB as well as in children are likely different from those with subjects with drug-susceptible TB and among adults. Therefore: 1. can the authors provide the exact number of such studies among the total 33 studies? i.e. (a) studies that include combination of DR-TB & DS-TB patients; and (b) studies including adults and children. 2. Did a separate evaluation of the prediction models from the studies specified above given consistent results to the remaining studies or are the findings similar?

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

1) Title: Please, make the title more concise by specifying 'drug-sensitive pulmonary tuberculosis in adults'.

Response: Thank you for this suggestion. We revised to the title to be “A systematic review of prediction models for pulmonary tuberculosis treatment outcomes in adults”. Though the focus of the review was drug-susceptible TB outcomes, we included some studies that included a small subset of individuals with MDR or other drug-resistance and, additionally, some studies did not report on drug-resistance. Because of this, we thought it would be misleading to include “drug-sensitive” in the title,

given the models evaluated in the review are not strictly all for drug-susceptible pulmonary TB patients.

2) Abstract: Please, also consider including the above specification in the abstract. The methods should also provide the method of data synthesis.

Response: We added additional details to clarify the study population of interest, and a sentence about how data were synthesized.

3) Introduction: The authors assert that 'to date, there has not been a formal synthesis or quality assessment of existing prediction models for TB treatment outcomes, which is essential to determine which models should inform clinical practice'. This statement is untrue. Van Wyk and colleagues reported findings of a systematic review of prediction models for prevalent pulmonary tuberculosis in 2017. It is strange that the authors did not pick up this study in their search. Because of this deficit, the authors have been unable to perform a critical appraisal of this important preliminary study in the introduction or a comparative analysis with their study in the discussion.

Response: Thank you for this important point. However, the referenced article focused solely on prediction models for prevalent TB – otherwise known as diagnostic prediction models – aimed to estimate the probability of having TB, among all patients in that specific setting, rather than the risk of unsuccessful TB treatment outcomes, among a population of TB patients. Because the focus of our review was on prognostic models (predicting treatment outcome, rather than prevalence), we did not include the information from the article mentioned. We clarified this further in the discussion section (1st paragraph).

4) Please, check the attached CHARMS checklist to ensure that it is completed.

Response: Thank you for mentioning this. The CHARMS checklist provides an outline for which data to extract from prognostic model studies, rather than a checklist for what is reported in systematic reviews of prognostic model studies. Thus, we present it as supplementary material to outline the data elements that were extracted in our review, but completing it would not be directly relevant for this study.

Reviewer: 2

1. Include your PROSPERO identification number for this systematic review.

Response: Thank you for this suggestion. The PROSPERO ID (CRD42020155782) has been added to the abstract and methods sections.

2. Is there any number of people who review the quality assessment?

Response: This is a good question, thank you. Data extraction and quality assessment were carried out simultaneously and by dual independent review. Two authors (LSP and FMR) independently extracted data and assessed the quality of included studies. We clarified this in the methods.

3. What are the review implications of your study?

Response: We believe the main implications of this study are that all prediction models for tuberculosis treatment outcomes have a high risk of bias in their analysis, and that future model development studies should better adhere to the recommended guidelines for developing prediction

models. The TRIPOD guidelines are available for exactly this purpose. Prediction models for TB treatment outcomes could be very informative for allocating resources or informing treatment approaches, but they must be rigorously developed and validated in order to be useful. We added additional details to the final paragraph of the discussion to address this point.

Reviewer: 3

In lines 48 - 51 (page 7), the authors stated that "articles that included drug-susceptible and drug resistant-cases or a combination of adult and children were included." The factors that influence the TB treatment outcomes, and thus prediction models, in subjects with drug-resistant TB as well as in children are likely different from those with subjects with drug-susceptible TB and among adults. Therefore:
1. can the authors provide the exact number of such studies among the total 33 studies? i.e. (a) studies that include combination of DR-TB & DS-TB patients; and (b) studies including adults and children.

Response: Thank you for this thoughtful question. Some of this information is available in Table 4, and we added additional detail to the manuscript to clarify (Results, study characteristics, 3rd paragraph). To answer your question directly, of the 33 studies included, 8 explicitly stated that they included MDR patients (and provided details on the percentage of the population that had MDR), whereas 7 studies explicitly excluded MDR patients, and 18 studies did not report on either specifically including or excluding persons with MDR TB. Twelve studies specifically included non-MDR drug resistance, 1 study explicitly excluded non-MDR forms of drug resistance, and 20 studies did not mention anything about non-MDR drug resistance. Regarding age, only 18 studies reported the age distribution of their population, and only 17 studies reported the minimum age included in their study. Of these, 7 (41%) had a minimum age of 15, 1 had a minimum age of 16 (6%), 1 had a minimum age of 17 (6%).

2. Did a separate evaluation of the prediction models from the studies specified above given consistent results to the remaining studies or are the findings similar?

Response: This is a great question as well. Unfortunately, because many studies did not report on the prevalence of drug resistance or on their age distribution, it is quite difficult to compare the results in these subgroups. Though the focus of our review was on drug-susceptible tuberculosis, it is possible that some studies that did not report on the prevalence of drug resistance in their study included only MDR patients, though we believe this is unlikely. In roughly comparing studies that include and exclude MDR patients, it seems the performance (c-statistic) for the studies that excluded MDR patients was slightly higher than studies that include MDR patients as a subset of their total study population (median: 0.78 vs. 0.74), but overall quality and other study characteristics seem relatively similar. When comparing studies that included younger adults vs. those with an age cut-off of 18 and above, the studies that included younger participants had somewhat better performance (median c-statistic: 0.82 vs. 0.74), but quality was similar. We did not include these results in the manuscript as we do not believe they add substantial value beyond the current text, within the journal constraints on our word count, and because their interpretation is limited by missing information in the majority of included studies.

VERSION 2 – REVIEW

REVIEWER	Benjamin Momo Kadia Liverpool School of Tropical Medicine, UK
REVIEW RETURNED	16-Jan-2021

GENERAL COMMENTS	Thank you for taking time to carry out the necessary revisions. The comments raised during the first round of peer review have been satisfactorily addressed. Please, consider revising the manuscript for any English or grammatical errors that may have been inadvertently missed.

REVIEWER	Azeez Adeboye University of Fort Hare South Africa
REVIEW RETURNED	13-Jan-2021

GENERAL COMMENTS	The authors have attended to the suggestions accordingly

REVIEWER	Dr. Toyin Togun, MD PhD London School of Hygiene and Tropical Medicine, London, United Kingdom. and MRC Unit The Gambia at the LSHTM Fajara, The Gambia.
REVIEW RETURNED	20-Jan-2021

GENERAL COMMENTS	"In roughly comparing studies that include and exclude MDR patients, it seems the performance (c-statistic) for the studies that excluded MDR patients was slightly higher than studies that include MDR patients as a subset of their total study population (median: 0.78 vs. 0.74), but overall quality and other study characteristics seem relatively similar. When comparing studies that included younger adults vs. those with an age cut-off of 18 and above, the studies that included younger participants had somewhat better performance (median c-statistic: 0.82 vs. 0.74), but quality was similar. We did not include these results in the manuscript as we do not believe they add substantial value beyond the current text." Given that the authors showed in their response to the initial critique that the performance of their prediction models might be influenced including or excluding MDR patients and by age, I do not agree that the decision not to include these results in the manuscript as it speaks to the overall performance of the prediction models. Given the concern about word count, they can include these result in the supplementary materials but important should also discuss these in the limitations of the study/model performance.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Dr. Benjamin Momo Kadia, Foumbot District Hospital, Grace Community Health and Development Association

Comments to the Author:

Thank you for taking time to carry out the necessary revisions. The comments raised during the first round of peer review have been satisfactorily addressed. Please, consider revising the manuscript for any English or grammatical errors that may have been inadvertently missed.

RESPONSE: Thank you for this suggestion. We have thoroughly reviewed the manuscript to revise grammatical errors and improve clarity.

Reviewer: 2

Dr. A Azeez, Department of Statistics, University of Fort Hare

Comments to the Author:

The authors have attended to the suggestions accordingly

Reviewer: 3

Dr. Toyin Togun, London School of Hygiene & Tropical Medicine

Comments to the Author:

"In roughly comparing studies that include and exclude MDR patients, it seems the performance (c-statistic) for the studies that excluded MDR patients was slightly higher than studies that include MDR patients as a subset of their total study population (median: 0.78 vs. 0.74), but overall quality and other study characteristics seem relatively similar. When comparing studies that included younger adults vs. those with an age cut-off of 18 and above, the studies that included younger participants had somewhat better performance (median c-statistic: 0.82 vs. 0.74), but quality was similar. We did not include these results in the manuscript as we do not believe they add substantial value beyond the current text."

Given that the authors showed in their response to the initial critique that the performance of their prediction models might be influenced including or excluding MDR patients and by age, I do not agree that the decision not to include these results in the manuscript as it speaks to the overall performance of the prediction models. Given the concern about word count, they can include these result in the supplementary materials but important should also discuss these in the limitations of the study/model performance.

RESPONSE: Great suggestion. We added a supplemental file to address this point (Supplemental File 8, Methods p. 10, Results p. 14). We further addressed the limitations of study population heterogeneities throughout the manuscript (Article Summary p.4, Discussion p. 15, 17-18).