

Prognostic accuracy of existing childhood asthma predictive models

To the Editor,

Systematic reviews of the growing literature on pre-school asthma prediction can help determine whether the available evidence supports the use of existing prognostic models to inform childhood asthma prevention and clinical management. The most salient prognostic review questions should address not only the accuracy of predictions derived from a prediction model but more importantly how the results affect patient management and outcomes.

As correctly noted by Kothalawala et al.¹ in their recent review “Prediction models for childhood asthma: A systematic review” “exemplary reviews of such prognostic models are not widely available. We praise the authors for their rigor of review methods used to summarize existing literature on school-age asthma prediction models. However, like in previous reviews,²⁻⁴ the distinction between diagnostic and prognostic tests remains unclear. By definition, a diagnostic test is used to help determine whether a patient has a disease of interest or not at the time the test is performed; its sensitivity (true positive proportion) and specificity (true negative proportion) is benchmarked against a ‘gold standard’ or reference test. In contrast, a prognostic model or test is used to predict a patient’s likelihood of developing a disease. As such, the reference standard for a prognostic test should be the observed proportion of the population who develop the disease of interest at a pre-defined future time-point.

In this article, we aim to corroborate the findings from the above systematic review with meta-analysis results, argue contrarily to some of the review’s conclusions and add to the recommendations for future studies.

In the abstract, the authors write “Overall model performance, assessed by area under the receiver operating curve (AUC), ranged between 0.66 and 0.87.”¹ Building on their findings, we performed a meta-analysis of the prediction parameters (sensitivity [*Se*] and specificity [*Sp*]: Figure 1); the pooled estimates from a bivariate random effects model (Figure 2) showed poor accuracy with a pooled *Se* of 0.51 (95%CI: 0.41, 0.61), *Sp* of 0.84 (95%CI: 0.77, 0.89) and an AUC of 0.72 (95%CI: 0.70, 0.75). Here, our pooled AUC estimate represents the overall performance of examined prognostic models at various diagnostic thresholds used to discriminate between future cases and non-cases of asthma. In fact, the partial AUC,⁵ which considers only those regions of the ROC space where data have been observed, or which correspond to clinically relevant values of *Se* and *Sp* is even lower (AUC: 0.66; 95%CI: 0.62, 0.70). It is also important to note that these statistic estimates do not account for the potential of reference standard misclassification error (since a gold standard diagnostic test for childhood asthma does not exist) which may further lower the magnitude and precision of these performance metrics.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/pai.13386](https://doi.org/10.1111/pai.13386)

This article is protected by copyright. All rights reserved

Bivariate-random effects meta-regression models (Table 1) used to explore the heterogeneity between potential moderators of study effect sizes (i.e. prognostic model accuracy) showed no statistically significant differences based on a child's genetic risk of asthma (e.g. parental history of asthma), use of weighted model scores, and prevalence of future asthma (Table 1). As expected, though with marginal statistical significance, externally validated models had higher *Se* ($b_{se}=0.83, p=0.06$) but lower *Sp* (higher false positive rate [b_{fpr}]: $0.94, p=0.05$) than non-validated models. The stringent Asthma Predictive Index (API) had lower *Se* ($b_{se}=-1.08, p=0.02$) but a higher *Sp* (lower false positive rate [b_{fpr}]: $-1.30, p=0.02$) than the loose API. Regression based models had a lower *Se* ($b_{se}=-3.32, p<0.01$) than machine-learning based models, but no statistically significant difference was observed for the false-positive rate ($b_{fpr}=0.46, p=0.53$). Related to the prognostic timeframes of existing models, we found a positive but not statistically significant relationship ($b_{se}=0.34, p=0.34$; $b_{fpr}=0.12, p=0.75$) between the accuracy of existing prognostic models and target age. This suggests that the performance of existing prognostic models may not be worse or better even when a child is older (target age range: 0-5 years). In addition, though not statistically significant, there was a weak negative relationship between prognostic accuracy and prediction age ($b_{se}=-0.05, p=0.64$; $b_{fpr}=-0.06, p=0.63$). Contrary to expectation, this suggests prognostic accuracy during early childhood may not be different than diagnostic accuracy at older ages (6-13 years) when diagnostic testing is expected to have higher accuracy. An alternative explanation for this null result is that time-dependent exposures probably modify disease pathogenesis and consequently alter future diagnostic outcomes as a child grows older.

Overall, however, because the predictive performance measures (AUC, sensitivity, specificity, positive predictive value, negative predictive value and likelihood ratios) of the prognostic models summarized by Kothalawala et al.¹ were determined retrospectively they are of limited clinical utility for the evaluation of childhood asthma risk prediction markers. These backward-looking measures do not summarize the predicted outcome probabilities and do not directly address questions about the impact of prognostic models on patient outcomes.⁶⁻⁸ Moreover, they also do not account for the fact that prognostic models accuracy may change over time as more children develop asthma, the outcome being predicted. Indeed, established childhood asthma risk factors⁹ such as parental asthma, prenatal environmental tobacco smoke and prematurity tend to have marginal impact on the discrimination statistics (e.g. AUC) individually yet lead to more accurate classification of large proportions of children into higher-risk or lower-risk categories. This is particularly true for children who tend to have persistent wheeze symptoms.¹⁰ Therefore, if the objective of childhood asthma prognostic model systematic reviews is to improve clinical decision-making, the precision and accuracy of estimates of asthma probabilities within different prognostic groups should be the primary focus.

For practical purposes, **first**, we suggest that it would be more useful to group the results of an asthma prognostic model into parsimonious categories corresponding to the implications for decision making. For example, if the actions that might follow a prognostic model are no further evaluation or treatment of “low” risk, further tests/monitoring for “intermediate” risk, or initiation of treatment or prevention in “high” risk children, then it would be useful to structure the review according to these prognostic categories (low, intermediate, and high risk) and clearly define each group including its outcome probabilities. This task can be implemented using reclassification tables⁷ to determine how a prognostic model affects the classification of infants/children into different prognostic groups to help inform the need for appropriate primary or secondary preventive interventions.

Second, it may also be informative to categorize infants or children using prognostic models as those who did or did not experience the predicted asthma outcomes within a specified time interval, and then look back to categorize the prognostic model’s risk strata and its accuracy for that point in time. Results from different studies with predicted outcome probabilities within prognostic groups at different time-points can then be meta-analyzed to validate the consistency or reliability of prognostic accuracy estimates.

Third, prognostic models should be used to predict beneficial or adverse responses to childhood asthma primary and secondary interventions. Evidence about the value of a predictive model typically should be presented as separate estimates of the intervention or treatment effect in subgroups defined by the predictive model, along with a statistical test for interaction between the treatment groups and subgroups defined by a predictive model to determine if the model’s predictive performance varies depending on prognostic risk group categorization.

Arthur H. Owora, MPH, DrPH¹; Yijia Zhang, MPH¹

Author Affiliations

¹Department of Epidemiology and Biostatistics, School of Public Health, Indiana University, Bloomington, Indiana. USA

References

1. Kothalawala DM, Kadalayil L, Weiss VBN, et al. Prediction models for childhood asthma: A systematic review. *Pediatric Allergy and Immunology*. 2020;31(6):616-627.
2. Smit HA, Pinart M, Antó JM, et al. Childhood asthma prediction models: a systematic review. *Lancet Respir Med*. 2015;3(12):973-984.

- Accepted Article
3. Luo G, Nkoy FL, Stone BL, Schmick D, Johnson MD. A systematic review of predictive models for asthma development in children. *BMC Med Inform Decis Mak*. 2015;15:99-99.
 4. Pinart M, Smit HA, Keil T, Bousquet J, Antó JM, Lødrup-Carlsen KC. Systematic review of childhood asthma prediction models. *European Respiratory Journal*. 2015;46(suppl 59):PA4513.
 5. Walter SD. The partial area under the summary ROC curve. *Stat Med*. 2005;24(13):2025-2040.
 6. Poses RM, Cebul RD, Collins M, Fager SS. The importance of disease prevalence in transporting clinical prediction rules. The case of streptococcal pharyngitis. *Annals of internal medicine*. 1986;105(4):586-591.
 7. Cook NR. Use and Misuse of the Receiver Operating Characteristic Curve in Risk Prediction. *Circulation*. 2007;115(7):928-935.
 8. Pepe MS, Janes HE. Gauging the performance of SNPs, biomarkers, and clinical factors for predicting risk of breast cancer. *Journal of the National Cancer Institute*. 2008;100(14):978-979.
 9. Castro-Rodriguez JA, Forno E, Rodriguez-Martinez CE, Celedón JC. Risk and Protective Factors for Childhood Asthma: What Is the Evidence? *J Allergy Clin Immunol Pract*. 2016;4(6):1111-1122.
 10. Owora AH, Zhang Y. Childhood wheeze trajectory-specific risk factors: a systematic review and meta-analysis. *Pediatric allergy and immunology : official publication of the European Society of Pediatric Allergy and Immunology*. 2020.

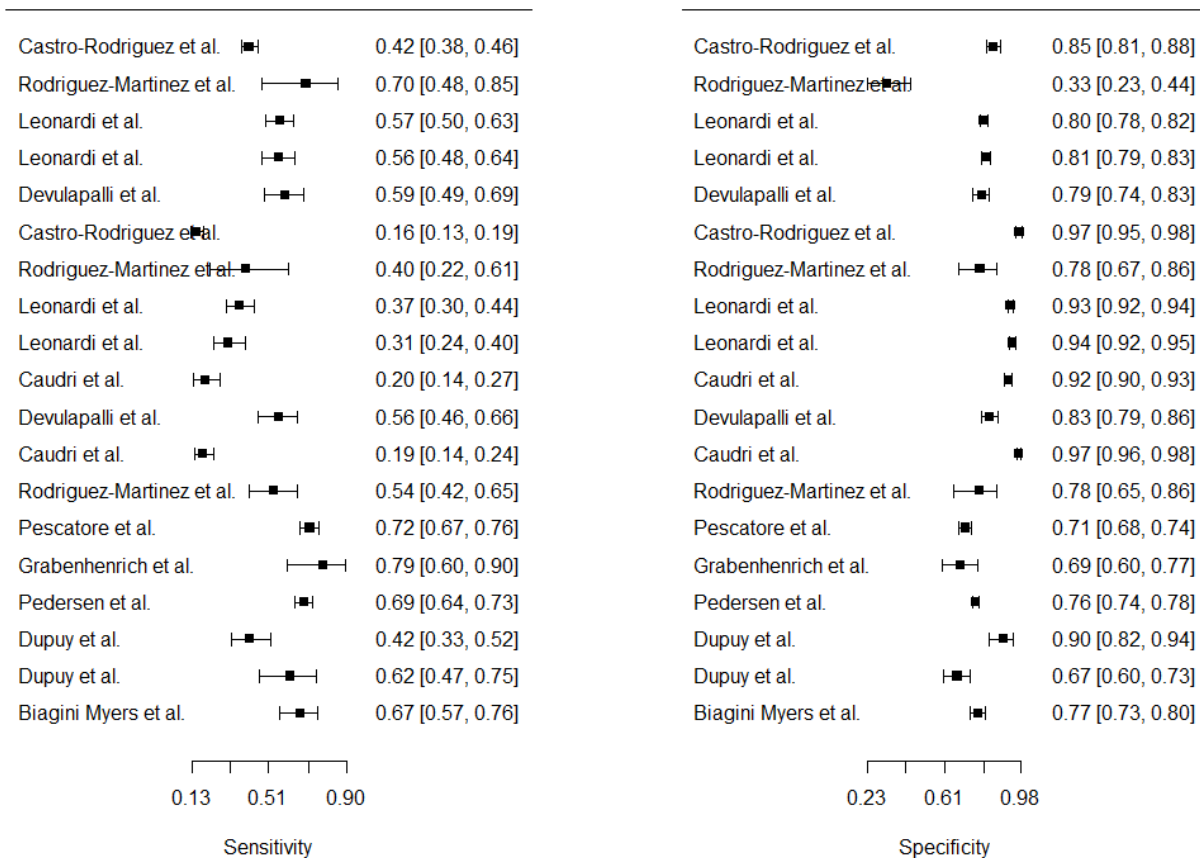


Figure 1. Forest plot of sensitivity and specificity estimates from existing childhood asthma prognostic models.

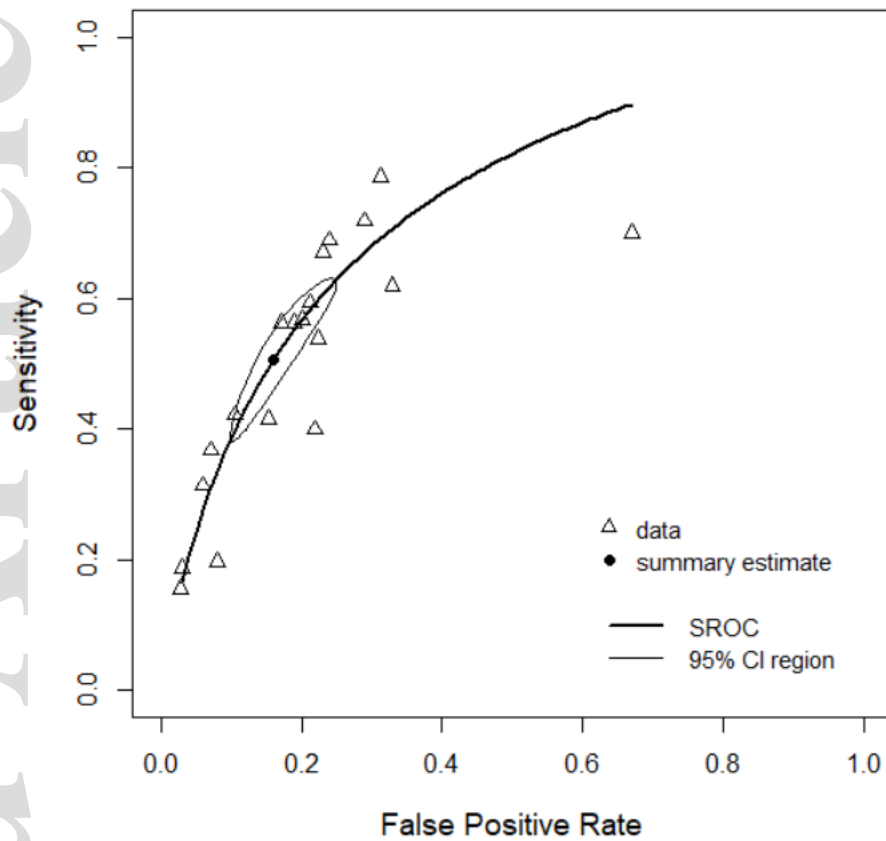


Figure 2. Summary Receiver Operating Curve (SROC) of childhood asthma prediction models

Table 1. Results from a bivariate meta-regression of the accuracy of childhood asthma prognostic models

Variable	Sensitivity Proportion		False Positive Proportion	
	<i>b</i> (SE)	<i>p</i> -value	<i>b</i> (SE)	<i>p</i> -value
Study population				
General population	Ref.	–		
At-risk	0.54 (0.41)	0.19	0.67(0.44)	0.13
Assessment tools				
Loose API	Ref.	–		
Stringent API	-1.08 (0.48)	0.02	-1.30(0.54)	0.02
Others	-0.05 (0.45)	0.91	-0.40(0.50)	0.43
External validation				
No	Ref.	–		
Yes	0.83 (0.44)	0.06	0.94(0.48)	0.05
Weighted scores				
No	Ref.	–		
Yes	0.54 (0.41)	0.19	0.30(0.47)	0.52
Statistical methods				
Machine learning	Ref.	–		
Regression-based	-3.32 (0.66)	<0.01	0.46(0.73)	0.53
Prevalence of asthma	-0.01 (0.01)	0.44	-0.01(0.01)	0.74
Target age	0.34 (0.35)	0.34	0.12(0.39)	0.75
Prediction age	-0.05 (0.11)	0.64	-0.06(0.12)	0.63

Statistically significant regression coefficients are in **bold**.