Treatment-decision algorithms for childhood pulmonary TB: Review of individual-patient data (IPD)

Kenneth S. Gunasekera,[1] Ted Cohen,[1] James A. Seddon[2,3]

[1] *Yale School of Public Health, New Haven, United States*
[2] *Imperial College London, London, United Kingdom*
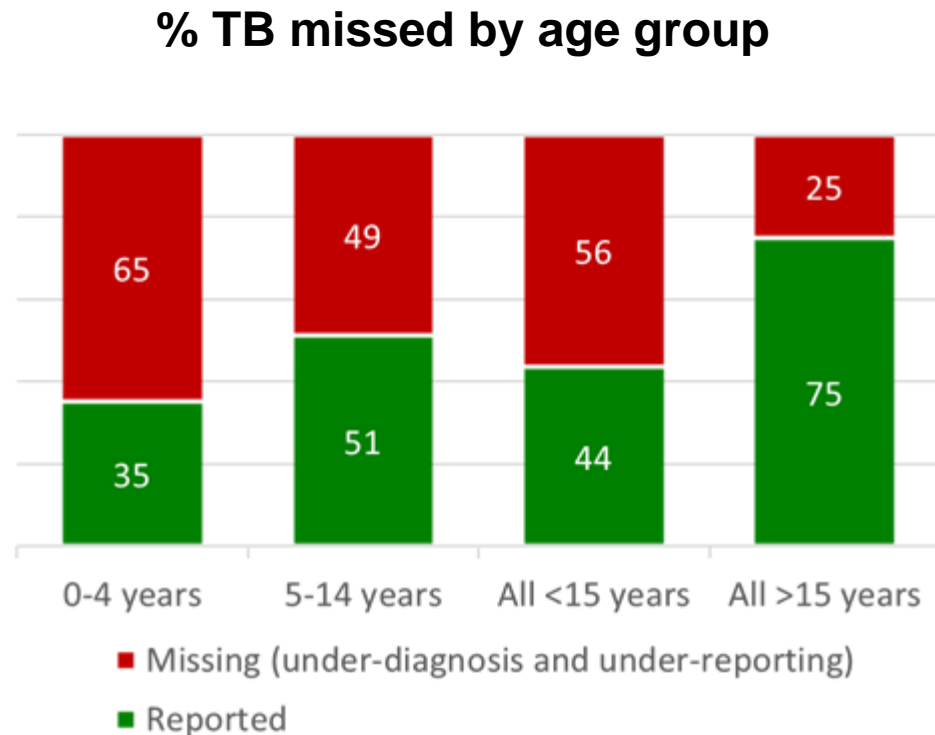[3] *Stellenbosch University, Cape Town, South Africa*

November 30, 2021

Annual meeting of the Child and Adolescent TB Working Group
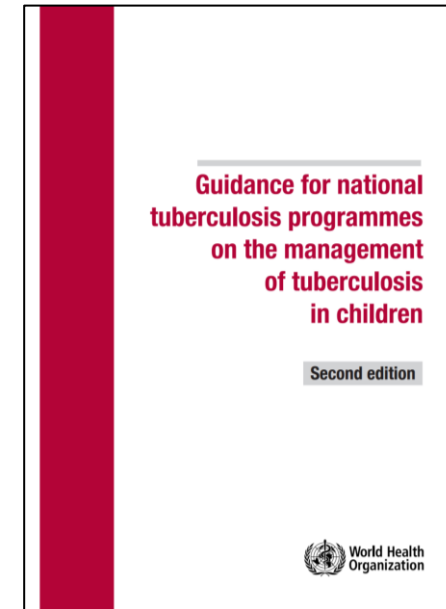
Disclosures & Conflicts of Interest

- Agreement for Performance of Work from WHO

- The reviewer team led development of the *Gunasekera et al., 2021* Algorithm that is evaluated

- The reviewer team has had scientific input from individuals involved in development of other algorithms being evaluated in this review (*Marcy et al., 2019*, *Marais et al., 2006*)

Diagnostic challenges for child pulmonary TB (PTB) contribute to child mortality

% TB missed by age group

Missing (under-diagnosis and under-reporting)
Reported

WHO 2020

Guidance for national tuberculosis programmes on the management of tuberculosis in children

Second edition
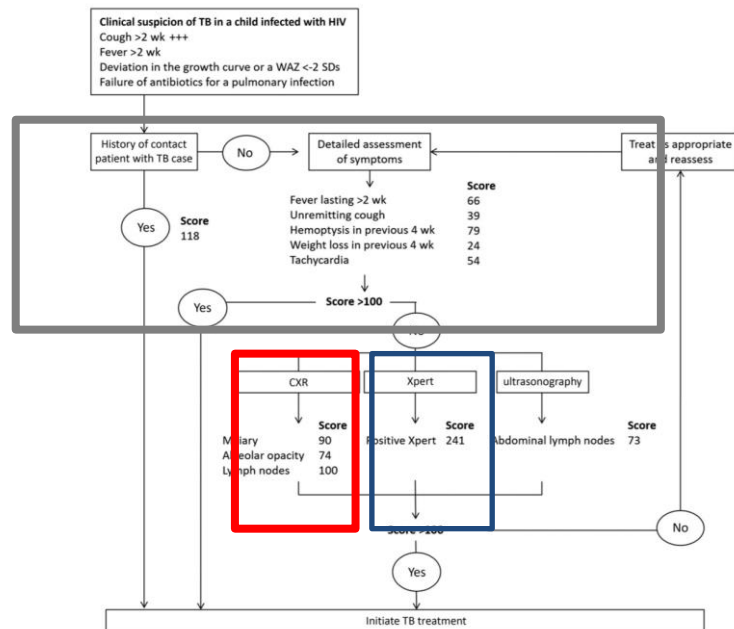
World Health Organization

Box 1. **Guidance on approach to diagnosis of TB in children**
- Careful history (including history of TB contact and symptoms consistent with TB)
- Clinical examination (including growth assessment)
- Tuberculin skin testing
- Chest X-ray (if available)
- Bacteriological confirmation whenever possible
- Investigations relevant for suspected pulmonary TB and suspected extrapulmonary TB
- HIV testing

Scores/algorithms standardize rapid treatment decision-making

Marcy et al. Pediatrics. 2019
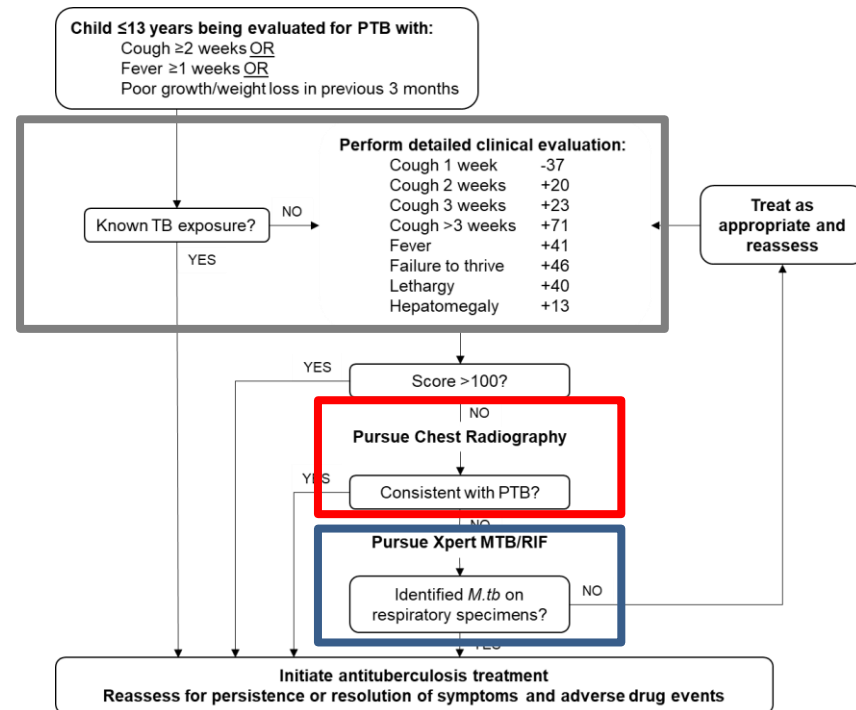
Gunasekera et al. Clin Infect Dis 2021

Clinical history/physical evaluation
Chest radiography
Bacteriology

ESTABLISH A LARGE, GEOGRAPHICALLY DIVERSE DIAGNOSTIC EVALUATIONS DATASET OF CHILDREN BEING EVALUATED FOR PTB

1. Evaluate existing scores/algorithms
2. Develop a data-driven algorithm

ASSEMBLE INDIVIDUAL PARTICIPANT DATA OF CHILDREN BEING EVALUATED FOR PTB

Data reflects population of children brought to healthcare with PTB symptoms

Demographic characteristics	Clinical evaluation	Diagnostic tests and imaging	Reference classification
Age Sex HIV-status Weight Height	Cough (duration) Fever (duration) Lethargy Weight loss Known contact w/ TB Temperature Heart rate Respiratory rate Etc.	Chest X-Ray *Features seen on chest X-ray* Rapid molecular test	TB (confirmed and unconfirmed) *OR* Unlikely TB

Total size:	4811
% TB:	38%
Age (months) median [IQR]:	26 [13.4-58.25]
% HIV-positive:	20%
% Severely acutely malnourished:	14%
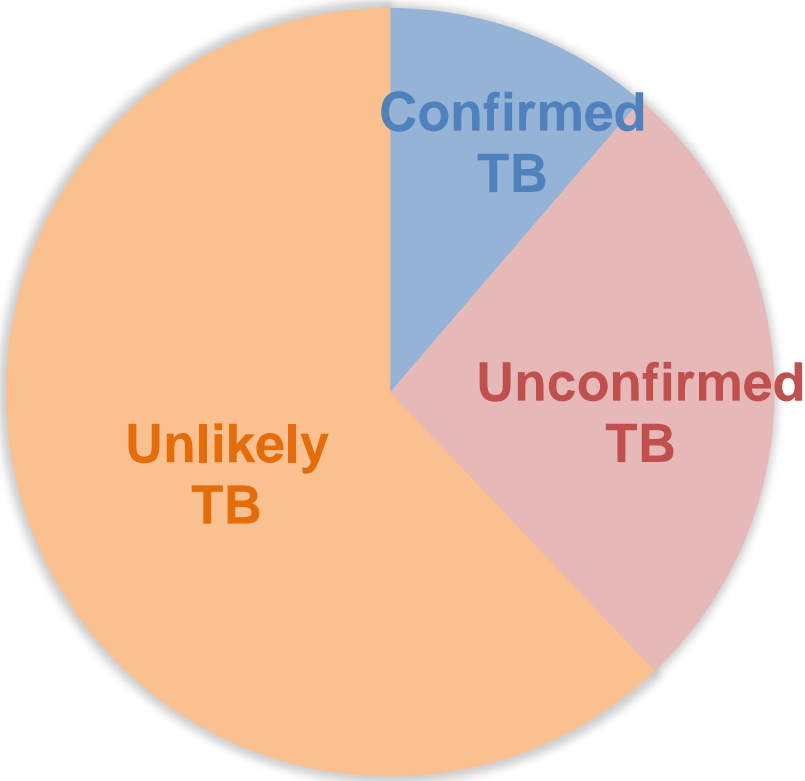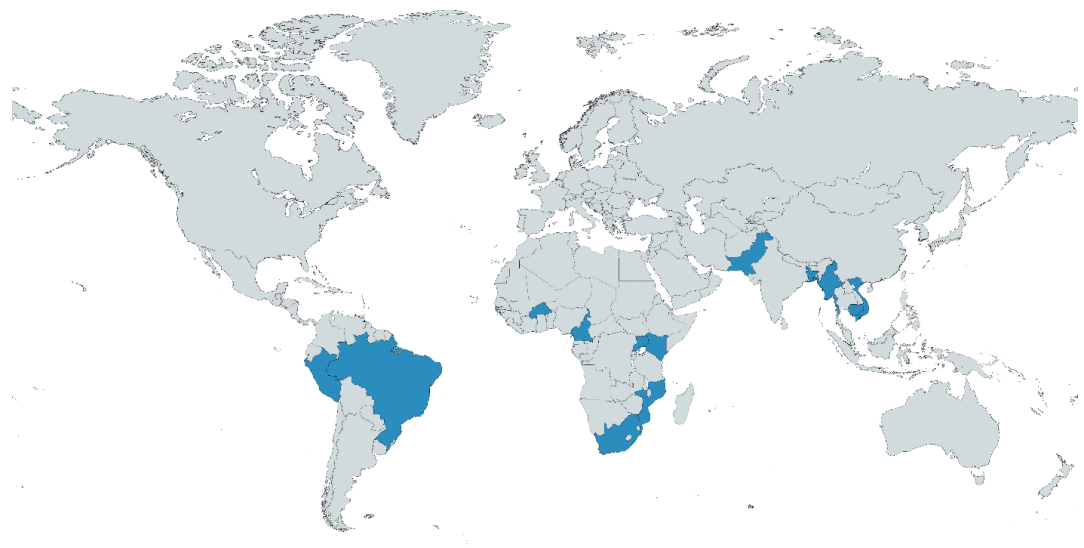
Reasonable attempts to handle imperfect data

Missing data

Study	Cough	CXR-nodes
1) A	1	1
2) A	NA	0
3) B	1	NA
4) B	0	NA
5) B	1	NA

MICE: Multiple Imputation by Chained Equations

Heterogeneous definitions

- i.e., weight loss:
  - Failure to thrive
  - Caregiver-reported weight loss
  - <-2 standard deviations below mean weight-for-age *Z*-score

Collapsed heterogeneous definitions where reasonable

1. EVALUATE EXISTING SCORES/ ALGORITHMS

Selected key algorithms/scores to evaluate ability to discriminate TB vs. non-TB

Algorithms
1) Union Desk Guide
2) Uganda National TB/Leprosy Control Program Algorithm
3) Brazilian Ministry of Health Child PTB Scoring System (cutoff of at least 30)
4) Gunasekera et al., 2021 Algorithm (HIV-negative children)
5) Keith Edward Score
6) Marcy et al., 2019 Algorithm (children living with HIV)
7) Stegen-Toledo Score (cutoff of at least 5)
8) Marais et al., 2006 Criteria

Modifications to scores/algorithms if IPD data not available

Example: Uganda National TB/Leprosy Control Program Algorithm
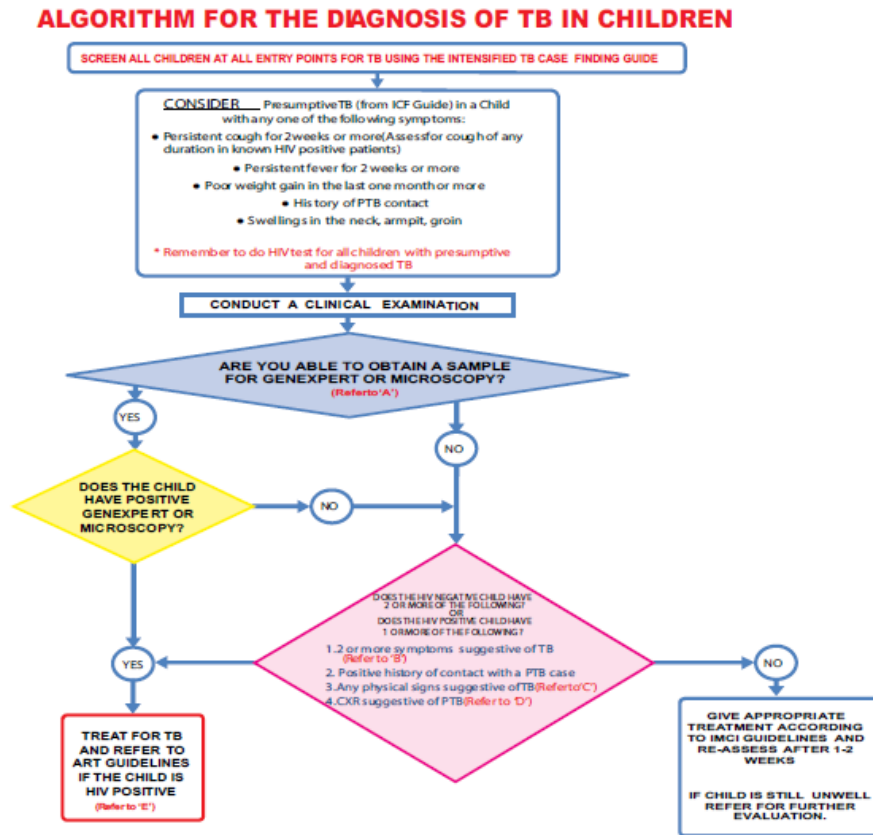
- Excluded
  - Antibiotic treatment history
  - Acute/recurrent pneumonia
  - Spinal deformity

Uganda Ministry of Health, Uganda National Tuberculosis and Leprosy Control Programme. *Manual for management and control of tuberculosis and leprosy in Uganda*. Kampala, Uganda: MoH, 2017.

Union Desk Guide: Standard-of-care
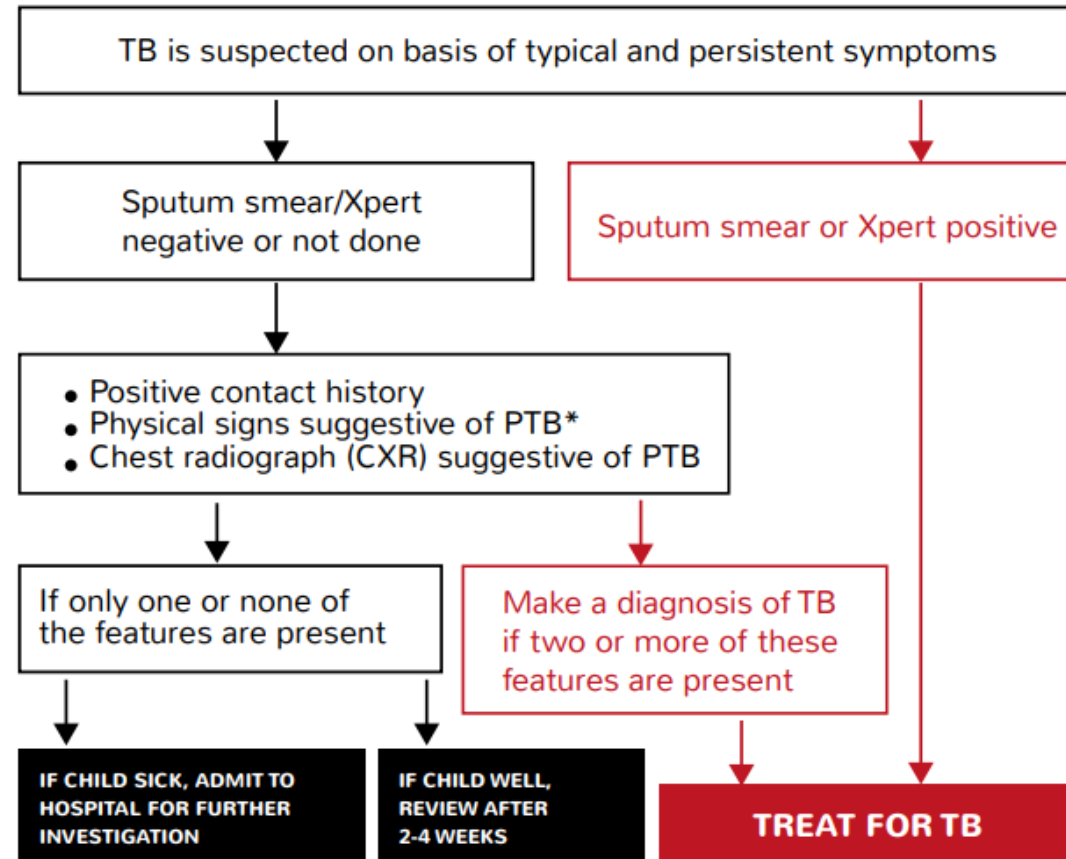
Performance against the Union Desk Guide is varied

Pooled Sensitivity Comparison

Union Desk Guide

Standard of Care	0.63 [0.51−0.74]
Uganda NTLP Algorithm	0.84 [0.74−0.91]
Stegen−Toledo Score	0.55 [0.39−0.7]
Marcy et al. Algorithm	0.69 [0.52−0.82]
Marais et al. Criteria	0.16 [0.09−0.27]
Keith Edward Score	0.92 [0.86−0.96]
Gunasekera et al. Algorithm	0.95 [0.88−0.98]
Brazilian MoH Score (cutoff 30)	0.86 [0.77−0.91]

Pooled Specificity Comparison

Union Desk Guide

|	
Standard of Care	0.65 [0.47−0.8]
Uganda NTLP Algorithm	0.39 [0.21−0.59]
Stegen−Toledo Score	0.85 [0.75−0.91]
Marcy et al. Algorithm	0.57 [0.41−0.72]
Marais et al. Criteria	0.89 [0.8−0.95]
Keith Edward Score	0.18 [0.1−0.31]
Gunasekera et al. Algorithm	0.09 [0.03−0.24]
Brazilian MoH Score (cutoff 30)	0.41 [0.19−0.67]

2. DEVELOP A DATA-DRIVEN ALGORITHM

1. Prediction modeling in algorithm development
2. Improve prediction in primary care/peripheral health centers

Prediction modeling in data-driven algorithm development

Clin Eval + Testing Model	
Feature	**Odds Ratio**
Cough Duration	0.62
Cough < 1 week	1.29
Cough 1-2 weeks	1.35
Cough 2-3 weeks	2.48
Cough > 3 weeks	
Fever	1.69
Failure to Thrive	1.80
Lethargy	1.68
History of TB Exposure	6.99
Hepatomegaly	1.18
CXR	9.38
Xpert	90.41
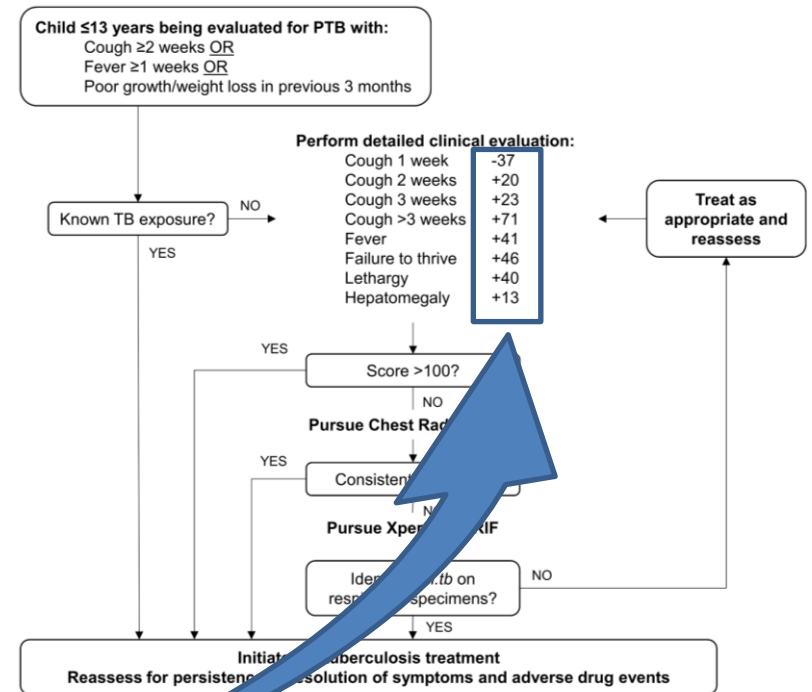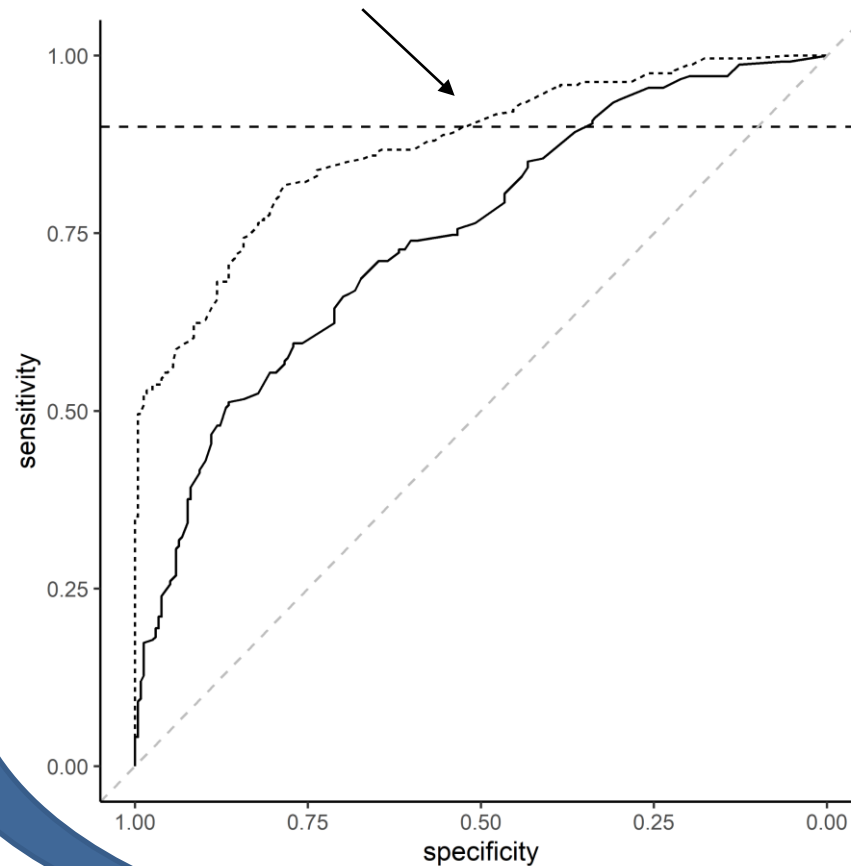
Scale odds ratio to score
>100 is TB at 90% sensitivity

Example from Gunasekera et al., Clin Infect Dis 2021

Considerations in selecting model sensitivity/specificity threshold

Sensitivity   Specificity
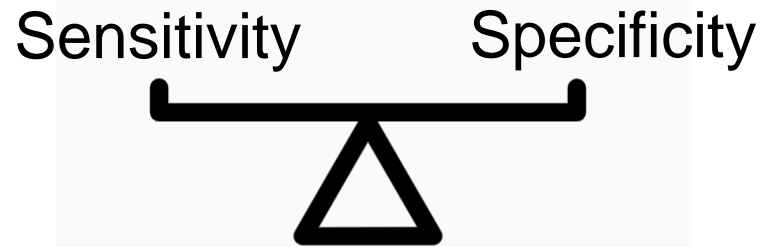
	Implications	
Decision	**Positive**	**Negative**
More sensitive threshold	• Reduce mortality due to TB disease progression	• Delayed/missed non-TB diagnosis • Unnecessary treatment
More specific threshold	• Pursue non-TB diagnosis	• Mortality due to TB disease progression • Lost to follow-up

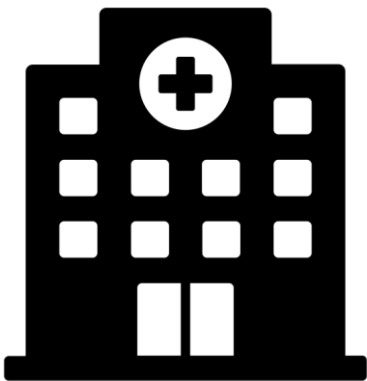# Differences between model development and model application populations
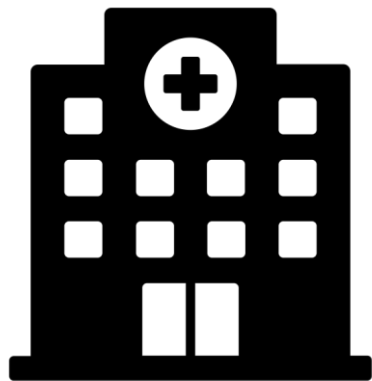
Model Development

**Tertiary/Referral
↑ TB Prevalence**

Model Application

**Tertiary/Referral
↑ TB Prevalence**

**Primary/Peripheral
↓ TB Prevalence**

Model performance expected to be consistent in high-TB prevalence, tertiary/referral care setting

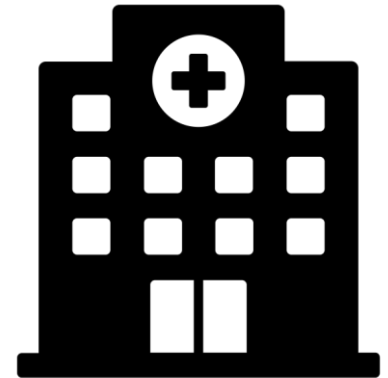**Model Development**

Model Application

Tertiary/Referral ↑ TB Prevalence

Tertiary/Referral ↑ TB Prevalence

Created by Adrien Coquet from Noun Project
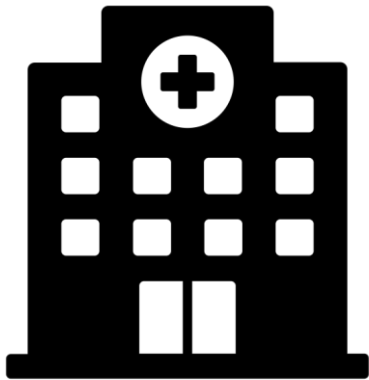
Created by Adrien Coquet from Noun Project

Created by IcoLabs from Noun Project

Primary/Peripheral ↓ TB Prevalence

Model performance may be worse in low-TB prevalence, primary/peripheral care setting

Model Development

Model Application

Tertiary/Referral ↑ TB Prevalence

Primary/Peripheral ↓ TB Prevalence

Tertiary/Referral ↑ TB Prevalence

Adding a triage step to delay treatment for children at low risk of TB-mortality may improve prediction

Triage

Child less than 10 years old with symptoms suggestive of pulmonary TB

Does the child have danger signs? — YES → Stabilize/refer to higher care

NO ↓

Is the child high-risk for TB mortality?

RETAIN / TRANSFER → Exit algorithm.

YES → Collect respiratory specimens for Xpert or LF-LAM and enter prediction model

NO → Treat non-TB, follow-up 1-2 weeks

Symptoms persist/worsen? — NO → Exit algorithm.

YES → Collect respiratory specimens for Xpert and enter prediction model

Assess for danger signs

Stratify by risk of progression of TB
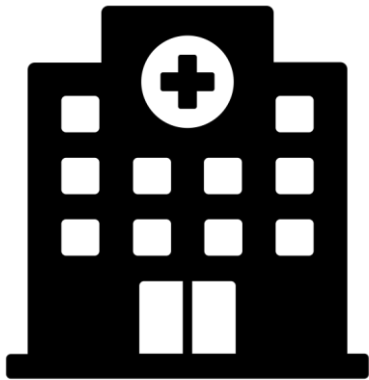- Higher risk, proceed
- **Lower risk, follow-up before entering model**

Prediction Model

Xpert or LF-LAM detect *Mycobacterium tuberculosis*?

YES / NO / NOT PERFORMED

Close contact w/ TB in previous 12 mos.?

YES / NO

Prediction model: clinical history/exam and chest x-ray

Sum scores of features present from clinical history/exam and chest x-ray

Do not treat for TB; follow-up in 1-2 weeks

Sum of scores > 100? — YES / NO

Prediction model to classify TB vs. non-TB

Initiate appropriate TB treatment

Need to be humble about algorithm sensitivity/specificity expectations on implementation

Model Development

Model Application

Child less than 10 years old with symptoms suggestive of pulmonary TB

?????

Tertiary/Referral
↑ TB Prevalence

Primary/Peripheral (w/ triage)
↑ TB Prevalence

Tertiary/Referral
↑ TB Prevalence

Limitations

1. Imperfect reference standard
2. Heterogeneous inclusion criteria, variable definitions
3. Missing data
4. No external validation

Strengths and future work

1. Evidence-based approach to treatment decision-making
2. Framework for future
   1. Better reference standard
   2. POC biomarkers

3. Future: External validation

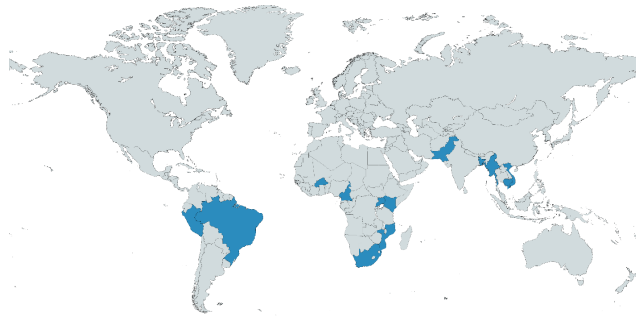# Acknowledgements

Yale SCHOOL OF PUBLIC HEALTH

Yale SCHOOL OF MEDICINE
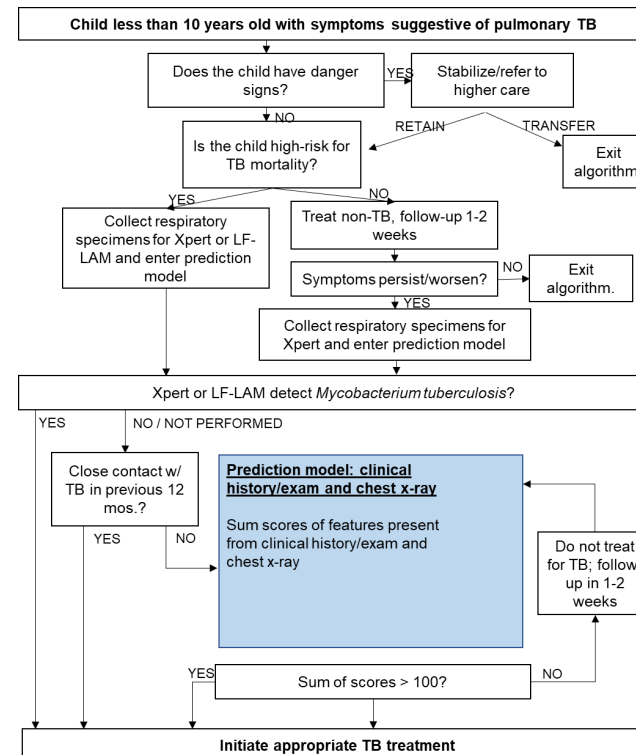
IPD treatment-decision algorithm development

Assembled large, geographically diverse cohort

Estimated the performance of existing scores/algorithms

Developing a prediction model to include in data-driven algorithm to guide childhood pulmonary TB treatment decision-making

Ken Gunasekera

@kennyguna

Ted Cohen

James Seddon

Total size:	4811
% TB:	38%
Age (months) median [IQR]:	26 [13.4-58.25]
% HIV-positive:	20%
% Severely acutely malnourished:	14%

Child less than 10 years old with symptoms suggestive of pulmonary TB

Does the child have danger signs? — YES → Stabilize/refer to higher care

NO

Is the child high-risk for TB mortality?

RETAIN / TRANSFER → Exit algorithm.

YES / NO

Collect respiratory specimens for Xpert or LF-LAM and enter prediction model

Treat non-TB, follow-up 1-2 weeks

Symptoms persist/worsen? — NO → Exit algorithm.

YES

Collect respiratory specimens for Xpert and enter prediction model

Xpert or LF-LAM detect *Mycobacterium tuberculosis*?

YES / NO / NOT PERFORMED

Close contact w/ TB in previous 12 mos.?

YES / NO

Prediction model: clinical history/exam and chest x-ray

Sum scores of features present from clinical history/exam and chest x-ray

Do not treat for TB; follow-up in 1-2 weeks

Sum of scores > 100? — YES / NO

Initiate appropriate TB treatment