2021

Prediction Tool to Identify Children at Highest Risk of Tuberculosis Disease Progression Among Those Exposed at Home

Meredith B. Brooks
Harvard Medical School

Leonid Lecca
Harvard Medical School

Carmen Contreras
Partners In Health/Socios En Salud

Roger Calderon
Universidade Federal do Rio de Janeiro

Rosa Yataco
Partners In Health/Socios En Salud

Open Forum Infectious Diseases

MAJOR ARTICLE

Prediction Tool to Identify Children at Highest Risk of Tuberculosis Disease Progression Among Those Exposed at Home

Meredith B. Brooks,[1] Leonid Lecca,[1,2] Carmen Contreras,[2] Roger Calderon,[2,3] Rosa Yataco,[2] Jerome Galea,[1,4,5] Chuan-Chin Huang,[1] Megan B. Murray,[1] and Mercedes C. Becerra[1,2]

[1]Department of Global Health and Social Medicine, Harvard Medical School, Boston, Massachusetts, USA, [2]Partners In Health/Socios En Salud, Lima, Peru, [3]Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, [4]School of Social Work, University of South Florida, Tampa, Florida, USA, and [5]College of Public Health, University of South Florida, Tampa, Florida, USA

Background. There is a dearth of research to understand which children, among those who are exposed at home to tuberculosis (TB), are at the highest risk of TB disease, to tailor care. We sought to identify predictors of TB progression in children.

Methods. We conducted a prospective cohort study of children living with adults with pulmonary TB in Lima, Peru (2009–2012). We applied classification and regression tree analysis to examine potential predictors of incident TB disease during 12 months in 3 age groups (0–4, 5–9, and 10–14 years). We calculated the relative risk (RR) for top predictors in each age group.

Results. Among 4545 children 0–14 years old, 156 (3.4%) were diagnosed with TB within 1 year of household exposure to TB (3.4%, 2.3%, and 4.7% in children 0–4, 5–9, and 10–14 years old, respectively). The most important predictor of TB was having a positive tuberculin skin test (TST) result, with RRs of 6.6 (95% CI, 4.0–10.7), 6.6 (95% CI, 3.2–13.6), and 5.2 (95% CI, 3.0–9.0) in the age groups 0–4, 5–9, and 10–14 years, respectively. In young children with a positive TST, not using isoniazid preventive treatment further increased risk of disease (RR, 12.2 [95% CI, 3.8–39.2]).

Conclusions. We present a tool that identifies child household contacts at high risk of TB disease progression based on data collected during contact tracing. In addition to the use of TB preventive therapy for all children exposed at home to TB, those children at highest risk of progressing to TB disease may benefit from more frequent follow-up.

Keywords. CART analysis; decision trees; pediatrics; tuberculosis.

Tuberculosis (TB) remains the leading infectious killer of adults in the world, with large populations facing high and stagnant case rates of this preventable disease [1]. An emerging network of coalitions—the Zero TB Initiative—seeks to rapidly drive down TB case rates in geographically defined zones, by locally deploying simultaneous strategies to increase case finding, access to treatment for all forms of TB disease, and access to TB preventive treatment (TPT) [2]. The use of targeted TPT—treatment that can stop TB infection from progressing to TB disease—is an essential component of a comprehensive strategy for TB elimination [3]. However, TPT has been vastly underutilized in areas where the TB burden is concentrated, remnants of a substandard approach advised for poorer countries for decades [2].

At the United Nations High-Level Meeting on TB in 2018, a minimal target was set to provide TPT to at least 30 million people by 2022 [4]. With >15 million children being exposed to TB every year because they share a household with an individual with infectious TB [5], the United Nations target is barely sufficient to provide TPT to half of children exposed to TB during that timeframe. With updated international guidance that now indicates TPT for household contacts of any age [6], TB programs will decisively expand local TPT access.

To optimize their reach and effectiveness at a larger scale, TB programs will likely need to apply tiered screening and support approaches even among groups at high risk of infection. Notably, while it is well established that TPT can benefit close contacts of all ages, the risk of disease progression is not uniform among children; between 5.2% and 7.6% of children exposed at home will become sick with TB within 2 years; of those children who test positive for TB infection, the incidence is even higher, ranging from 8.8% to 19.0% [7]. Can those specific children—at far higher risk of disease progression than the others—be identified early? This would allow the design and refinement of more

Received 26 July 2021; editorial decision 20 September 2021; accepted 22 September 2021.

Correspondence: Mercedes C. Becerra, ScD, Department of Global Health and Social Medicine, Harvard Medical School, 641 Huntington Ave, Boston, MA 02115, USA (mercedes_becerra@hms.harvard.edu).

Open Forum Infectious Diseases®2021

https://doi.org/10.1093/ofid/ofab487

targeted longitudinal support, such as more frequent testing, to avert preventable disease, disability, and deaths.

A previous study identified an increased risk of incident TB disease in young children who had a conversion from negative to positive on sequential interferon-γ release assays, which detect *Mycobacterium tuberculosis* infection [8]. Another study identified that a quantified TB exposure score—incorporating various components related to a child's intensity of exposure to an index TB patient—was associated with increased risk of prevalent TB disease in child household contacts [9]. However, these studies were association-based and did not focus on identifying accurate classifications of disease status for individual children, as a predictive study would do. TB disease prediction tools are typically devised using clinical epidemiology methods, but rarely focus on children, especially those exposed to TB at home. Thus, among children who were exposed at home to TB, we sought to identify which children were at highest risk of disease progression, using machine-learning tools and data collected during TB household contact tracing.

PATIENTS AND METHODS

Study Design

We conducted a secondary analysis of data collected in a prospective cohort study of household contacts of pulmonary TB cases in Lima, Peru, between September 2009 and August 2012 [10]. Household members, of any age, of adults diagnosed with incident pulmonary TB at 106 participating health centers were enrolled. Household members were tested for TB infection using tuberculin skin tests (TSTs), and those reporting symptoms of TB disease at the time of enrollment were referred to the local health clinic for clinical evaluation. Those without TB disease at baseline were followed at 2, 6, and 12 months, at which time they were reevaluated for TB disease. Those without TB disease at the follow-up visits underwent repeat TST at 6 and 12 months. Routine care, as per National TB Program Guidelines, included the use of a BCG vaccine in all newborns and the provision of TPT for any individual up to age 19 years who lives with a TB patient [11]. We asked household members if they had been offered and initiated TPT. Those diagnosed with TB disease were referred to local health centers for treatment according to national guidelines. In this substudy, we focus only on household members who were <15 years of age at baseline and had TB disease ruled out.

Predictors and Outcome Variables

We assessed 21 potential predictors of TB disease in children aged 0–14 years who did not have TB disease at baseline. These variables included sociodemographic information, clinical characteristics, medical history, household characteristics, and test results collected at baseline. To ensure the predictors would be applicable in clinical practice, we assessed only those variables that could be readily collected at the time a child is evaluated for TB disease at a health facility.

The outcome of interest was the occurrence of incident TB disease during 12 months of follow-up, as determined by a positive test result using microbiological tests (smear or culture) or a clinical diagnosis, according to consensus guidelines for diagnosing TB disease in children [12].

Individual-level characteristics included age, sex, weight (kg), height (cm), body mass index (BMI; kg/m^2), malnutrition (defined as a sex-specific BMI-for-age *z* score <2, calculated per World Health Organization tables [13]), relation to the initial TB case (child, sibling, or other), socioeconomic status (classified as lowest, middle, and highest tertile from scores derived using principal components analysis of housing asset, weighted by household size [14]), human immunodeficiency virus (HIV) seropositivity, cardiac disease, asthma, number of BCG vaccination scars, history of TB disease, baseline use of isoniazid for TPT, and TST result at baseline (national guidelines defined a positive TST by an induration with a diameter of ≥10 mm). Household-level characteristics included the number of individuals living in the household and the type of housing (house, apartment, other).

Statistical Analysis

We reported participant characteristics in aggregate, by age group (0–4, 5–9, 10–14 years), and for individuals who developed TB disease. We used classification and regression tree (CART) analysis to search through all potential predictors and their possible cutoff values to identify the most important predictors and their optimum predictive thresholds, to distinguish between children who did and did not develop incident TB disease. We ranked and selected the primary node and assessed the relevance of each variable in the final model. Variable importance measures, as determined by computing the improvement measure attributable to each variable in its role as a surrogate to the primary split, were assigned to each potential predictor and entailed both marginal and interaction effects involving each variable. The data set was then split into increasingly homogenous subgroups, using improvement in the Gini gain score to split nodes and add smaller daughter nodes to the tree. At each daughter node, the CART algorithm selected the explanatory variable and splitting value that gave the best discrimination between the 2 outcome classes. Maximal trees were generated by splitting each daughter node until each outcome class was homogeneous or contained few observations. Then, trees were manually pruned based on relative misclassification costs, complexity, and parsimony.

Next, we applied receiver operating characteristic (ROC) analysis and performed 10-fold cross-validation, in which the whole data set was randomly split into learning and test data sets. CART analysis was applied to determine model performance and predictive accuracy in these test sets, removing the

need for a validation data set. We reported the final derived decision tree for each of the 3 age group models along with ROC area under the curve (AUC) for the full tree from the test set, and the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) and associated 95% confidence intervals (CIs).

To assess the results in more familiar terms to the clinical practitioners who would be the main users of the derived trees, we also report standard frequentist statistical estimates. For the final derived trees for each age group model, we assessed the utility of the predictors and identified thresholds by examining each node's crude association with developing TB disease using regression analysis. We report relative risks (RRs) and 95% CIs. Comparisons of the proportion experiencing the outcome of interest between these groups were made using χ^2 tests.

CART analysis was run using Salford Systems Data Mining and Predictive Analytics Software version 8.0 (Salford Systems, San Diego, California). Standard statistical analyses were performed using SAS version 9.4 software (SAS Institute, Cary, North Carolina).

Patient Consent Statement

A parent or guardian was required to provide written informed consent for all children participating in the household cohort study. In addition, written informed assent to participate was obtained for children 8 years of age and older; it was determined that children <8 years of age could participate with only their parent or guardians written informed consent. The household cohort study was approved by the Institutional Review Board of the Harvard School of Public Health (Boston, Massachusetts) and the Research Ethics Committee of the National Institute of Health of Peru (Lima, Peru).

RESULTS

The parent study enrolled 4500 initial TB cases and 14 044 of their household members, of whom 4623 were 0–14 years of age. A total of 78 child household members were excluded because they had TB disease at baseline. Of the remaining 4545 children, 1768 (38.9%) were 0–4 years of age, 1452 (31.9%) were 5–9 years, and 1325 (29.2%) were 10–14 years. One-quarter (24.1%) of all children had a positive TST at baseline, with the percentage increasing with age (0–4 years: 16.6%; 5–9 years: 25.6%; 10–14 years: 32.2%). A total of 156 (3.4%) children were diagnosed with TB disease during follow-up, with 3.4%, 2.3%, and 4.7% among the age groups 0–4, 5–9, and 10–14 years, respectively. Table 1 summarizes the demographic and clinical characteristics for all children in aggregate, by age group, and for those who developed TB.

We evaluated 17 potential predictors for the age group model 0–4 years (Figure 1). The primary node was the child's TST result; among children aged 0–4 years who had a positive TST result, 11.9% developed TB disease compared to only 1.8% who had a negative TST result (RR, 6.6 [95% CI, 4.0–10.7]; $P < .001$). In children with a positive TST result at baseline, those who did not use TPT had a higher risk of developing TB disease during the year of follow-up (24.0%) than those who did use TPT (2.0%) (RR, 12.2 [95% CI, 3.8–39.2]; $P < .001$). In children with a negative TST result at baseline, those living in a household with ≤3 other individuals had a higher risk of TB disease (5.4%) than those with >3 (1.4%) (RR, 3.9 [95% CI, 1.8–8.6]; $P = .002$). The discriminatory properties of the model were sensitivity of 50.0% (95% CI, 36.8%–63.2%), specificity of 83.2% (95% CI, 81.3%–84.9%), PPV of 9.5% (95% CI, 7.4%–12.1%), and NPV of 97.9% (95% CI, 97.4%–98.4%). The AUC for the test set was 0.693.

We evaluated 20 potential predictors for the age group model 5–9 years (Figure 2). The primary node was the child's TST result; among children aged 5–9 years with a positive TST result, 6.2% developed TB disease, as compared to 0.9% (RR, 6.6 [95% CI, 3.2–13.6]; $P < .001$) with a negative TST result. The discriminatory properties of the model were sensitivity of 70.6% (95% CI, 52.5%–84.9%), specificity of 74.3% (95% CI, 72.0%–76.6%), PPV of 6.2% (95% CI, 5.0%–7.7%), and NPV of 99.1% (95% CI, 98.4%–99.4%). The AUC in the test set was 0.673.

We evaluated 20 potential predictors for the age group model 10–14 years (Figure 3). The primary node was the child's TST result; among children aged 10–14 years who had a positive TST result, 10.1% developed TB disease, as compared to 1.9% (RR, 5.2 [95% CI, 3.0–9.0]; $P < .001$) with a negative TST result. In children who had a positive TST result at baseline, 11.0% of those whose weight was <62.1 kg developed TB, whereas no child whose weight was >62.1 kg did. The discriminatory properties of the model were as follows: sensitivity of 67.7% (95% CI, 54.7%–79.1%), specificity of 69.2% (95% CI, 66.6%–71.7%), PPV of 9.7% (95% CI, 8.2%–11.6%), and NPV of 97.8% (95% CI, 96.8%–98.4%). The AUC was 0.660 in the test set.

DISCUSSION

Across all age groups, having a positive TST was the most important predictor of incident TB disease during a 1-year follow-up period. The estimated risk was 5–6 times that of children who tested negative for TB infection. This finding is consistent with existing literature, including a recent meta-analysis that identified the 2-year cumulative risk of TB in children exposed at home to TB who tested positive for TB infection as 8.8%–19.0%, which was substantially higher than when TB infection status was not taken into account (5.2%–7.6%) [7]. While the AUCs for each model indicate relatively low discrimination overall, the NPVs were all >97%, making the derived trees potentially attractive screening tools for identifying children who will develop incident TB over a year of follow-up. The tools can be used to rule out children at low risk of incident TB disease

Table 1. Baseline Characteristics of Children Exposed at Home to Tuberculosis

Characteristics	Age Group				
	All Children (N = 4545)	0–4 y (n = 1768)	5–9 y (n = 1452)	10–14 y (n = 1325)	Children Who Developed TB Disease (n = 156)
Age, y, mean (SD)	6.5 (4.3)	2.0 (1.4)	6.9 (1.4)	12.0 (1.4)	7.2 (4.5)
Female sex	2252 (49.6)	877 (49.6)	713 (49.1)	662 (50.0)	83 (53.2)
Weight, kg, median (IQR)	22.8 (14.5–37.0)	13.0 (10.3–16.0)	25.0 (20.6–30.5)	44.7 (36.5–52.0)	21.0 (15.1–39.7)
Height, cm, median (IQR)	116 (95–138)	89 (77–99)	120 (113–128)	147 (140–154)	115 (96–144)
BMI, kg/m^2, median (IQR)	17.7 (16.0–20.1)	16.8 (15.5–18.4)	17.2 (15.6–19.4)	20.0 (18.0–22.6)	17.7 (16.1–19.8)
Relation to initial TB patient	n = 4542	n = 1766	n = 1451	n = 1325	n = 156
Child	1600 (35.2)	740 (41.9)	499 (34.4)	361 (27.3)	82 (52.6)
Sibling	692 (15.2)	94 (5.3)	217 (14.9)	381 (28.8)	27 (17.3)
Other	2250 (49.5)	932 (52.8)	735 (50.7)	583 (44.0)	47 (30.1)
Socioeconomic status	n = 4415	n = 1719	n = 1407	n = 1289	n = 147
Lower tertile	1710 (38.7)	666 (38.7)	571 (40.6)	473 (36.7)	65 (44.2)
Middle tertile	1961 (44.4)	777 (45.2)	613 (43.6)	571 (44.3)	60 (40.8)
Higher tertile	744 (16.9)	276 (16.1)	223 (15.9)	245 (19.0)	22 (15.0)
Household type	n = 4545	n = 1768	n = 1452	n = 1325	n = 156
House	3242 (71.3)	1222 (69.1)	1054 (72.6)	966 (72.9)	102 (65.4)
Apartment	768 (16.9)	336 (19.0)	228 (15.7)	204 (15.4)	33 (21.2)
Other	535 (11.8)	210 (11.9)	170 (11.7)	155 (11.7)	21 (13.5)
No. of individuals in the home, median (IQR)	6 (5–9)	6 (4–9)	6 (5–9)	6 (5–9)	6 (4–8)
Malnourished	n = 4500 158 (3.5)	n = 1750 91 (5.2)	n = 1436 48 (3.3)	n = 1314 19 (1.5)	n = 156 4 (2.6)
No. of BCG vaccine scars, median (range)	1 (0–3)	1 (0–3)	1 (0–2)	1 (0–3)	1 (0–2)
TST positive at baseline	n = 4360 1051 (24.1)	n = 1673 278 (16.6)	n = 1408 361 (25.6)	n = 1279 412 (32.2)	n = 150 99 (66.0)
TPT use	n = 4544 2300 (50.6)	n = 1767 910 (51.5)	n = 1452 766 (52.8)	n = 1325 624 (47.1)	n = 156 46 (29.5)
Prior TB disease	n = 4543 41 (0.9)	n = 1767 5 (0.3)	n = 1451 12 (0.8)	n = 1325 24 (1.8)	n = 156 2 (1.3)
Asthma	n = 4534 298 (6.6)	n = 1763 71 (4.0)	n = 1450 102 (7.0)	n = 1321 125 (9.5)	n = 156 10 (6.4)
Cardiac disease	n = 4541 62 (1.4)	n = 1767 23 (1.3)	n = 1451 14 (1.0)	n = 1323 25 (1.9)	n = 156 4 (2.6)
HIV positive	n = 4490 2 (0.04)	n = 1746 0 (0)	n = 1435 2 (0.1)	n = 1309 0 (0)	n = 154 0 (0)
Diagnosed with TB	156 (3.4)	60 (3.4)	34 (2.3)	62 (4.7)	NA

Data are presented as No. (%) unless otherwise indicated. Frequencies for smoking, alcohol use, diabetes, kidney disease, and high blood pressure not reported due to low or no occurrence in age groups.

Abbreviations: BMI, body mass index; HIV, human immunodeficiency virus; IQR, interquartile range; NA, not applicable; SD, standard deviation; TB, tuberculosis; TPT, tuberculosis preventive treatment; TST, tuberculin skin test.

and inform which group of children may benefit from more frequent follow-up and screening for disease, in order to capture cases of TB disease as promptly as possible.

TPT was identified as an important predictor in the age group 0–4 years; in the subgroup that had TB infection, not using TPT increased the risk of incident TB disease 12-fold. This is consistent with existing literature, in which, among children exposed to TB, those <1 year old who tested positive for TB infection and did not receive TPT had an 18% risk of incident TB disease over 2 years of follow-up; similarly, children 2–5 years of age had a risk of 19% [7]. Additionally, in children of all ages, TPT is found to be effective in 63% of all child contacts and in 91% of those with TB infection [7]. Although the use of TPT was not identified as an important predictor for the age group models 5–9 or 10–14 years, that does not indicate that TPT failed to prevent progression to TB disease in these groups; CART analysis uses recursive partitioning and selects important variables based on how well they discriminate between the levels of the outcome variable [15–17]. Thus, the established marked protection of incident TB through use of TPT should still be considered as essential for children of all ages.

Among children aged 0–4 years who had a negative TST result at baseline, those living in a household with ≤3 other individuals had a 4-fold higher risk of TB disease than those with >3 individuals in the household. This may be indicative of younger children relying more heavily on their caregiver

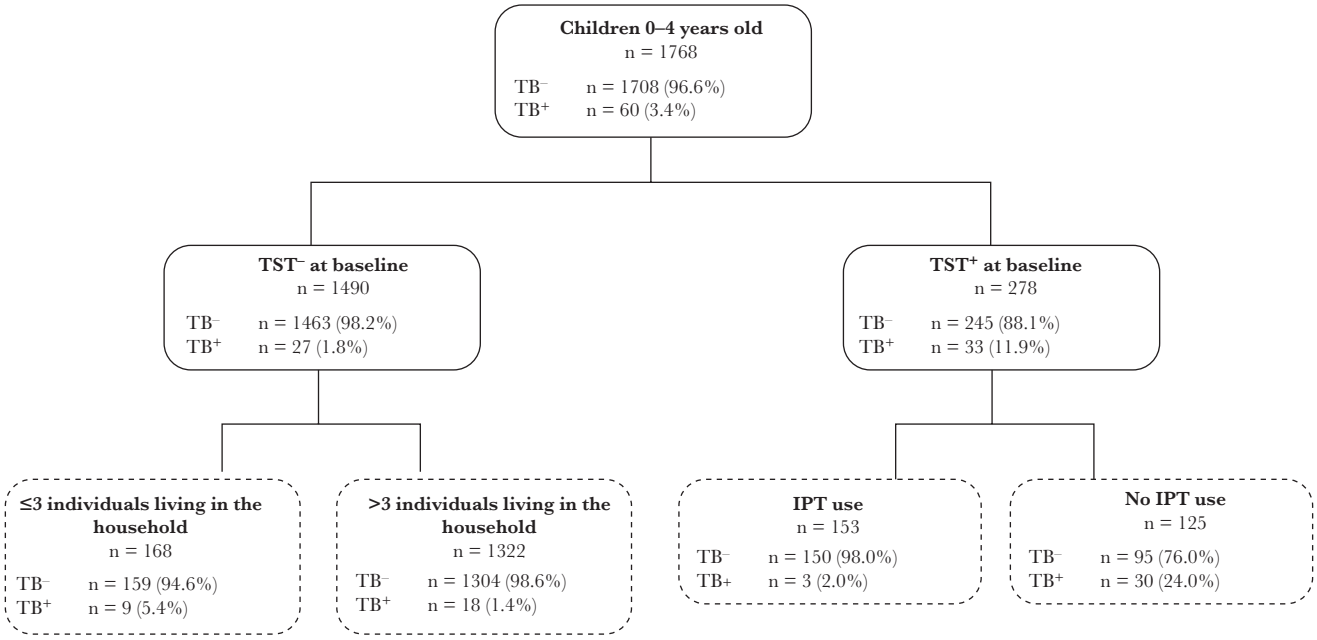

Figure 1. Classification and regression tree–derived predictors of tuberculosis disease in children aged 0–4 years (n = 1768). Terminal nodes are depicted by dashed lines. Abbreviations: ⁻, negative test result; ⁺, positive test result; IPT, isoniazid preventive therapy; TB, tuberculosis; TST, tuberculin skin test.

for daily activities and, with fewer adults in the household to serve as an alternative caregiver, still having intense and prolonged exposure with their primary caregiver who has TB disease. Additionally, we found that in children 10–14 years old who had a positive TST result at baseline, being below a certain weight threshold led to a much higher risk of TB disease than those above that threshold. This is consistent with the known increased risk of TB disease among individuals who are malnourished [18].

While there are other existing clinical prediction tools for children being evaluated for TB disease, this tool differs from those in several ways. First, these age-specific tools were derived using information about children who were household contacts of patients being treated for active TB, so they were all exposed at home and are at increased risk of infection and disease. This is in contrast to other tools that were developed to inform more rapid TB treatment-decision making in children presenting to health facilities with presumed TB disease [19, 20]. Second, our tools aim to predict incident TB disease over a year of follow-up to inform the frequency of and strategy for disease screening and tailored support, while other tools aim to more accurately detect TB disease at a cross-sectional snapshot: the time of presentation to the health facility [19, 20]. These tools were derived using information from children with HIV [20] and from HIV-negative children in South Africa [19]. Clinical prediction tools also exist in the context of Peru. Indeed, one utilized the same

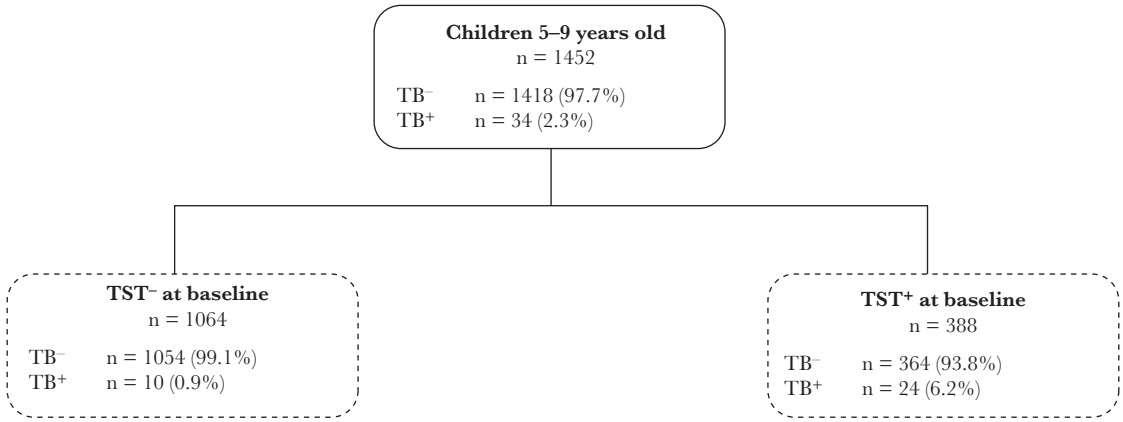

Figure 2. Classification and regression tree–derived predictors of tuberculosis disease in children aged 5–9 years (n = 1452). Terminal nodes are depicted by dashed lines. Abbreviations: ⁻, negative test result; ⁺, positive test result; TB, tuberculosis; TST, tuberculin skin test.

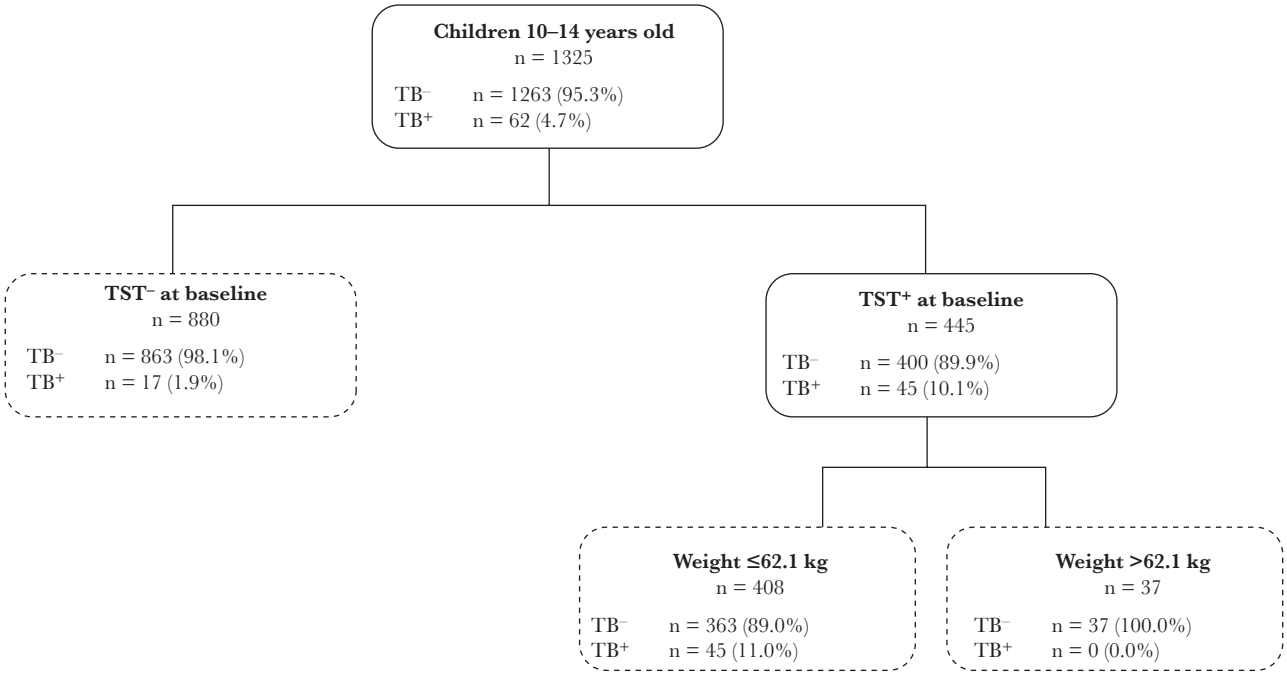

Figure 3. Classification and regression tree–derived predictors of tuberculosis disease in children aged 10–14 years (n = 1325). Terminal nodes are depicted by dashed lines. Abbreviations: ⁻, negative test result; ⁺, positive test result; TB, tuberculosis; TST, tuberculin skin test.

data set but focused only on adult contacts of TB patients [21]. The other focused on the identification of full households—not individuals—at high risk of TB disease [22]. The tools we produce in this report are a unique contribution to the literature because they are (1) derived from children exposed to TB at home; (2) predict TB disease up to a year after exposure; (3) are focused on children from a low-HIV prevalence setting; and (4) are age-specific.

A major strength of this approach is the use of CART analysis, which results in an easily interpretable decision tree that is ready for applicability into clinical practice with minimal training [15, 17]. CART is a useful exploratory analysis that can illuminate previously concealed links among important predictors and outcomes [23, 24]. While CART does not provide a statistical output such as a CI by which to quantify or support the validity of the findings, the results can be subjected to hypothesis testing using more standard statistical methods [25].

This study has several limitations. First, the parent study was limited to individuals >15 years old who had pulmonary, bacteriologically confirmed TB disease and their household members. This limits the generalizability of the models, as all children included were exposed to an adult who met these criteria. Second, there were other sources of data collected from household members, including the index TB patient, that were purposefully not included in the final models so that the final derived trees included only variables that were typically collected when household contacts are evaluated. By limiting the variables used, we ensured that all necessary data to apply the tool could be readily available or easy to collect during contact tracing. Third, although all children in this study had a household TB exposure, it is impossible to know whether they were also exposed and infected outside of the household. While the risk of infection and disease may be associated with the background rate of TB in the study area, that information is not readily available to incorporate into the analyses.

CONCLUSIONS

The derivation of decision trees to predict incident TB disease among children exposed at home to TB provides a tool that is simple to use and interpret and that can be readily applied for clinical application. The use of only data collected during contact tracing for potential inclusion within the decision trees further adds to their ease of use in a busy clinic setting. While all children exposed at home to TB disease are at high risk of progressing to TB disease, the use of these prediction tools may identify a subgroup of children at even higher risk than others. This may inform tailored support approaches and frequency of follow-up testing for these children to diagnose TB disease as early as possible.

Notes

Disclaimer. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding institutions.

Financial support. The original study was supported by the National Institutes of Health (NIH) (grant numbers U19 A1076217 and U01 AI057786 to M. B. M. and M. C. B.). Research reported in this publication was supported by the Thrasher Research Fund (Early Career Award number

15020); the American Lung Association (grant number CA-683831); and the National Institute of Allergy and Infectious Diseases of the NIH (award number K01AI151083 to M. B. B.).

Potential conflicts of interest. All authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. World Health Organization. Global Tuberculosis Control Report 2020. Geneva, Switzerland: WHO; **2020**.
2. Keshavjee S, Nicholson T, Khan AJ, Ditiu L, Farmer PE, Becerra MC. Tuberculosis epidemic control: a comprehensive strategy to drive down tuberculosis. In: Friedman L DM, Davies PDO. Clinical Tuberculosis. 6th ed. Boca Raton, FL: CRC Press.
3. Rangaka MX, Cavalcante SC, Marais BJ, et al. Controlling the seedbeds of tuberculosis: diagnosis and treatment of tuberculosis infection. Lancet **2015**; 386:2344–53.
4. Stop TB Partnership, United Nations Office for Project Services. United Nations High Level Meeting on TB Key Targets and Commitments for 2020. Geneva, Switzerland: United Nations; **2018**.
5. Dodd PJ, Gardiner E, Coghlan R, Seddon JA. Burden of childhood tuberculosis in 22 high-burden countries: a mathematical modelling study. Lancet Glob Health **2014**; 2:e453–9.
6. World Health Organization. WHO Guidelines on Tuberculosis Infection Prevention and Control. Geneva, Switzerland: WHO; **2019**.
7. Martinez L, Cords O, Horsburgh CR, Andrews JR; Pediatric TB Contact Studies Consortium. The risk of tuberculosis in children after close exposure: a systematic review and individual-participant meta-analysis. Lancet **2020**; 395:973–84.
8. Andrews JR, Nemes E, Tameris M, et al. Serial QuantiFERON testing and tuberculosis disease risk among young children: an observational cohort study. Lancet Respir Med **2017**; 5:282–90.
9. Malik AA, Amanullah F, Siddiqui S, et al. Quantified TB exposure can predict prevalent TB disease in child contacts. Pediatr Infect Dis J **2021**; 40:e208–9.
10. Becerra MC, Huang CC, Lecca L, et al. Transmissibility and potential for disease progression of drug resistant *Mycobacterium tuberculosis*: prospective cohort study. BMJ **2019**; 367:l5894.
11. Ministerio de Salud del Perú. Dirección General de Salud de las Personas: Estrategia sanitaria nacional de prevención y control de la tuberculosis norma técnica de salud para el control de la tuberculosis. **2006**. https://cdn.www.gob.pe/uploads/document/file/342511/Norma_t%C3%A9cnica_de_salud_para_el_control_de_la_tuberculosis20190716-19467-rmxgh7.pdf. Accessed 12 August 2021.
12. Graham SM, Ahmed T, Amanullah F, et al. Evaluation of tuberculosis diagnostics in children: 1. Proposed clinical case definitions for classification of intrathoracic tuberculosis disease. Consensus from an expert panel. J Infect Dis **2012**; 205(Suppl 2):S199–208.
13. World Health Organization. Child growth standards. **2011**. http://www.who.int/childgrowth/software/en/. Accessed 10 August 2020.
14. Filmer D, Pritchett LH. Estimating wealth effects without expenditure data—or tears: an application to educational enrollments in states of India. Demography **2001**; 38:115–32.
15. Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and Regression Trees. Boca Raton, FL: Chapman and Hall/CRC; **1984**.
16. Kim H, Loh W-Y. Classification trees with unbiased multiway splits. J Am Stat Assoc **2001**; 88:457–67.
17. Steinberg D, Colla P. CART: Tree-Structured Non-parametric Data Analysis. San Diego, CA: Salford Systems; **1995**.
18. Jaganath D, Mupere E. Childhood tuberculosis and malnutrition. J Infect Dis **2012**; 206:1809–15.
19. Gunasekera KS, Walters E, van der Zalm MM, et al. Development of a treatment-decision algorithm for HIV-uninfected children evaluated for pulmonary tuberculosis. Clin Infect Dis **2021**; 73:e904–12.
20. Marcy O, Borand L, Ung V, et al. A treatment-decision score for HIV-infected children with suspected tuberculosis. Pediatrics **2019**; 144:e20182065.
21. Li R, Nordio F, Huang CC, et al. Two clinical prediction tools to improve tuberculosis contact investigation. Clin Infect Dis **2020**; 71:e338–50.
22. Saunders MJ, Wingfield T, Datta S, et al. A household-level score to predict the risk of tuberculosis among contacts of patients with tuberculosis: a derivation and external validation prospective cohort study. Lancet Infect Dis **2020**; 20:110–22.
23. Campbell DK. Nonlinear science: from paradigms to practicalities. Los Alamos Sci Spec Issue **1987**; 218–62.
24. Campbell D, Farmer D, Crutchfield J, Jen E. Experimental mathematics: the role of computation in nonlinear science. Commun ACM **1985**; 28:374–384.
25. Kuhn L, Page K, Ward J, Worrall-Carter L. The process and utility of classification and regression tree methodology in nursing research. J Adv Nurs **2014**; 70:1276–86.