Clinical Infectious Diseases

MAJOR ARTICLE

Body Mass Index and Incident Tuberculosis in Close Tuberculosis Contacts

María B. Arriaga, PhD[1,2*], Gustavo Amorim, PhD [3*], Marina C. Figueiredo, MS[1], Cody Staats, MS[1], Afrânio L. Kritski, PhD[4], Marcelo Cordeiro-Santo, PhD[5,6,7], Valeria C. Rolla, PhD[8], Peter F. Rebeiro, PhD[1,3,9^], Bruno B. Andrade, PhD [10,11,12,13,14^], Timothy R. Sterling, MD[1^], for the RePORT-Brazil consortium[#]

[1]Division of Infectious Diseases, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA; [2]Fundação José Silveira, Centro de Pesquisa, Aprendizagem e Inovação, Salvador, Bahia, Brazil; [3]Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN, USA; [4]Programa Acadêmico de Tuberculose, Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil; [5]Fundação Medicina Tropical Doutor Heitor Vieira Dourado, Manaus, Brazil; [6]Programa de Pós-Graduação em Medicina Tropical, Universidade do Estado do Amazonas, Manaus, Brazil; [7]Universidade

[#]**Additional authors from the Regional Prospective Observational Research in TB (RePORT) Brazil consortium (corporate authorship):** Mariana Araújo-Pereira, Alice M. S. Andrade, Dayse Lira, Michael S. Rocha, Vanessa Nascimento, Pedro Brito, Saulo R. N. Santos, André Ramos, Alexandra Brito Souza, Renata Spener-Gomes, Amanda França Aguiar, Alysson G. Costa, Jaquelane Silva, Jamile G. de Oliveira, Aline Benjamin, Adriano Gomes-Silva, Flavia M. Sant'Anna, Francine P. Ignácio, Maria Cristina Lourenço, Adriana S. R. Moreira, Megan Turner, Solange Cavalcante, Betina Durovni

*MBA and GA contributed equally to this work.

^contributed equally to this work

Corresponding author: PhD. María B. Arriaga, Postdoctoral fellowship at Division of Infectious Diseases, Department of Medicine, Vanderbilt University School of Medicine: mbag711@gmail.com; Phone: +1 (629) 2548390, Nashville, Tennessee, USA.

Co-corresponding author: Dr. Timothy Sterling, Professor at Division of Infectious Diseases, Department of Medicine, Vanderbilt University School of Medicine: timothy.sterling@vumc.org; Phone: +1 (615) 3362309, Nashville, Tennessee, USA.

DOI: 10.1093/cid/ciaf475

Nilton Lins, Manaus, Brazil; [8]Instituto Nacional de Infectologia Evandro Chagas, Fiocruz, Rio de Janeiro, Brazil; [9]Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA; [10]Laboratório de Pesquisa Clínica e Translacional, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil; [11]Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Salvador, Brazil; [12]Curso de Medicina, Escola Bahiana de Medicina e Saúde Pública (EBMSP), Salvador, Brazil; [13]Instituto de Pesquisa Clínica e Translacional, Faculdade Zarns, Clariens Educação, Salvador, Brazil; [14]Curso de Medicina, Universidade Salvador, Salvador, Brazil

Background: Approximately 95% of people infected with *M. tuberculosis* do not progress to tuberculosis (TB) disease. Identifying key determinants of TB progression could focus prevention efforts.

Methods: Contacts of pulmonary TB patients were enrolled in a prospective multi-center cohort study (RePORT-Brazil) from 2015-2019 and followed for 24 months. Empirical review and LASSO regression, using baseline clinical and laboratory information, were used as dimension reduction techniques to determine factors for inclusion in prediction models. Models were created for: 1) all contacts, 2) contacts IGRA-positive at baseline, and 3) IGRA-positive contacts who did not receive TB preventive therapy (TPT; <30 days isoniazid). Internal validation was performed using bootstrapping.

Results: Among 1846 contacts of 619 TB index patients, 25 (1.4%) progressed to TB. No TPT was a risk factor for progression to TB among all contacts [mixed-effects adjusted hazard ratio (aHR)=16.55, 95% confidence interval (CI): 2.22-124.45]. Internal validation with all contacts yielded an area under the ROC curve of 0.80 [95%CI: 0.72-0.86]. Body mass index (BMI) was inversely associated with increased risk of progressing to active TB among IGRA-positive contacts who did not receive TPT (aHR=0.89, 95%CI: 0.80-0.99). IGRA-positive contacts with BMI <25 kg/m^2 had a 4.14-fold (95%CI: 1.17-14.67) higher risk of progression to TB than IGRA-positive contacts with BMI ≥25 kg/m^2: 8.4% vs. 2.1%, respectively.

Conclusions: BMI <25 kg/m^2, a readily available biomarker, identified IGRA-positive close TB contacts at high risk of progressing to TB disease. Prioritizing this high-risk group for TB preventive therapy could improve TB prevention efforts.

Key words: close contact; TB progressor; active TB; prediction model; tuberculosis; interferon gamma release assay; body mass index

INTRODUCTION

Although tuberculosis (TB) can be prevented, treated, and cured, it continues to be a major global public health problem; in 2023 there were an estimated 10.8 million incident TB cases and 1.25 million deaths in persons diagnosed with TB. [1]

Mycobacterium tuberculosis (Mtb) is transmitted via the respiratory route, from people with TB to close contacts who share the same space. Exposure to TB can lead to Mtb infection [2]. Approximately one-quarter of the global population has Mtb infection [3]. The lifetime risk of progressing from Mtb infection to TB disease is 5-10%, with approximately half of those progressing ≤2 years after infection [4].

Identifying the determinants of progression to TB among contacts is critical for TB control; prevention efforts could then focus on the 5-10% who would develop TB without intervention. Several studies have identified close contact characteristics associated with progression to TB: age (<5 years) [5], malnutrition or low body mass index (BMI) [6-8], HIV infection [9], and evidence of Mtb infection based on tuberculin skin test [10] or interferon-gamma release assay (IGRA)[11]. Conversely, TB preventive therapy (TPT) decreases the risk [10, 12]. There are also TB source case characteristics associated with progression to TB in their close contacts: positive sputum smear [13], increased age [14], male sex [15], and cavities on chest x-ray [15]. Nevertheless, these findings have varied between populations, and most predominant risk factors remain unclear. The ability to identify contacts at high risk of progressing to TB disease would facilitate an emphasis on TPT in these persons and thereby decrease the TB burden[2].

Previously, our group described the cascade of care among adult and child TB close contacts in Brazil [16] and associations between transcriptomic signatures in whole blood and progression to active TB [17]. To extend this research, the present study developed a prediction model to identify risk factors for progressing to active TB among close contacts of pulmonary TB index patients.

METHODS

Study Cohort and Design

This longitudinal observational study was conducted in the Regional Prospective Observational Research in Tuberculosis (RePORT)-Brazil cohort of culture-confirmed pulmonary TB patients and their close contacts. Participants were enrolled between June 2015 and June 2019 and followed for 24 months. Enrollment sites were in three Brazilian states (and four cities): Rio de Janeiro (Rio de Janeiro and Caxias), Bahia (Salvador), and Amazonas (Manaus). The representativeness of the RePORT-Brazil cohort compared to all TB patients reported in Brazil has been described previously [18].

Participants and study groups

Individuals ≥18 years old diagnosed with new or recurrent pulmonary TB (with or without extrapulmonary TB) by culture-positive sputum (Lowenstein-Jensen medium or BD BACTEC MGIT) were enrolled. Close contacts of enrolled TB index patients, defined as individuals exposed to a culture-positive pulmonary TB patient for ≥4 hours in one week in the 6 months prior to TB diagnosis, were identified, invited to be interviewed, and examined at RePORT-Brazil healthcare units. More information is provided in **Supplementary Figure 1**.

Contacts were classified by baseline IGRA result and TPT status: (i) all contacts, regardless of IGRA or TPT status; (ii) contacts who were IGRA-positive at baseline, and (iii) contacts who were IGRA-positive at baseline and did not receive TPT or received TPT for <30 days. These groups differed clinically but were not mutually exclusive.

Ethics statement

The RePORT-Brazil protocol, informed consent, and study documents were approved by the institutional review boards at each study site and at Vanderbilt University Medical Center.

Procedures

Enrolled contacts were investigated for Mtb infection and evaluated in-person at baseline and 6 months after enrollment; subsequent assessments every 6 months were by telephone, then in-person if they reported signs or symptoms of TB disease. At baseline, we performed a clinical evaluation; contacts were enrolled if they reported no TB symptoms, including weight loss. Chest X-ray was performed and blood collected for IGRA and HIV testing. Clinical, laboratory, and socio-demographic data were collected through standardized case report forms in REDCap[19]. All procedures were performed according to Brazilian National TB Guidelines[20]. Contacts with a negative or indeterminate baseline IGRA test underwent repeat testing at month 6. IGRA collection, processing and interpretation were performed according to the manufacturer's recommendations for the QuantiFERON®-TB Gold in tube assay (Qiagen) until May 2017, and the QuantiFERON® -TB Gold Plus assay subsequently.

Contacts i) with a positive IGRA result, or ii) <5 years old, or iii) living with HIV (ii and iii regardless of IGRA result), were recommended to receive TPT according to Brazilian guidelines[20]. The TPT regimen was isoniazid 5-10mg/kg daily in adults and 10 mg/kg in children (300 mg daily maximum for both) for 6-9 months. Complete TPT was defined as receiving ≥6 months of isoniazid.

Outcome

Progression to TB was defined as TB among contacts confirmed a) microbiologically by a positive Xpert (MTB/RIF or Ultra) or solid (Lowenstein-Jensen) or liquid (BD MGIT) culture; or b) clinically, based on abnormal chest X-ray, suggestive histology (e.g., caseating or necrotizing

granulomas) and clinical symptoms such as cough, fever, night sweats, and weight loss occurring within 24 months of enrollment.

Statistical analysis

Continuous variables were compared using the non-parametric Mann-Whitney *U* test (between two groups) or Kruskal-Wallis test (between more than two groups). Categorical variables were compared using Fisher's exact test (cell count <5) or Pearson's chi-square test (cell count ≥5). We assumed non-informative censoring for participants who did not complete the 24 months of follow-up. Only baseline data were used for analysis. P-values <0.05 were considered statistically significant.

Dimension reduction

Due to the low number of TB progressors among contacts and the high number of potential predictors, we performed dimension reduction in the following order: (i) empirical review of potential predictors and (ii) least absolute shrinkage and selection operator (LASSO) regression analysis [21, 22]. Only variables included in stage (i) were included in stage (ii).

A set of predictors was selected *a priori* and identified by consensus of co-authors. The empirical review excluded 1) similar variables (e.g., CAGE assessment and alcohol variables) that could be adequately explained or measured by other variable(s); 2) variables with the same value or absent in all progressors to TB. The LASSO regression further selected the variables to be included in the final prediction model. We used five-fold cross-validation to select the regularization parameter. Results using 10-fold and leave-one-out cross validation are available in the **Supplementary Appendix**.

We also conducted a secondary approach in which principal component analysis (PCA) was used as an intermediate step between the empirical review and the time-to-event LASSO regression [23] (see **Supplementary Methods**). PCA was conducted to remove correlated variables by clustering predictors; components were interpreted based on moderate correlation ($|r| \geq 0.6$).

Prediction models and validation

Three prediction models for time to TB, one for each group of interest, were performed using Cox models with LASSO [24]. Participants who did not progress to TB were censored at their last visit date or the end of the study period. The partial log-likelihood was used as a loss function for the time-to-event LASSO analysis [24].

We performed internal validation to assess model discrimination between progressors and non-progressors. Discrimination was quantified with the area under the receiver operating characteristic (ROC) curve; 200 bootstrapped samples were used to derive an optimism-corrected measure [25]. More details are in the **Supplementary Methods**.

Statistical inference

Variables selected for each study group after dimension reduction were entered into individual, mixed effects Cox models. We used a random intercept and treated index patients as clusters to account for possible correlation within close contacts of the same index patient. We assumed that each contact was related to a single index patient. Results are presented as adjusted hazard ratios (aHR) with corresponding 95% confidence intervals (CI). The proportional hazards assumption was evaluated using Schoenfeld residuals. The post-selection inference procedure, refitting the model after variable selection, does not account for potential variability introduced during variable selection [26]; p-values should therefore be interpreted with caution (see **Supplementary Methods**).

Participants with incomplete observations

We performed two analyses to account for missing data. A complete-case analysis, in which participants with missing values for variables retained after empirical review, were excluded from subsequent steps (e.g., excluded from the LASSO). Missingness was <1%, except for income and abnormal chest X-ray (missing for 3.4% and 10.8% of participants, respectively). We also re-ran analyses with an imputed dataset. Missing data were imputed using chained equations [27]. We used a single imputation, as missingness was generally low. Results using imputed data are presented in the manuscript, and the complete-case results are presented in the **Supplementary Appendix**.

Critical values for clinical interest and decision-making

The third model focused on IGRA-positive contacts at baseline who received no TPT or received TPT for <30 days. Due to limited number of events, the relationship between BMI and the risk of progressing to active TB was visually assessed using plots. The effect of BMI on TB progression alone was further explored, using logistic regression and restricted cubic splines with three knots. Although the fitted model suggested a non-linear relationship, it was less informative for clinical decision-making. Thus, we conducted a segmented (piecewise) logistic regression to estimate the critical value of BMI associated with a significant change in TB progression. The optimal 'breakpoint', i.e., the point where there was a significant change in the relationship between BMI and the outcome, was estimated via maximum likelihood.

RESULTS

Study population characteristics

The study population included 1846 close contacts of 619 pulmonary TB patients enrolled in RePORT-Brazil (**Supplementary Figure 1**), with 86% completing ≥1.5 years of follow-up. Among all close contacts, 678 were IGRA-positive, 1132 were IGRA-negative, and 31 were

IGRA-indeterminate at baseline. Five contacts did not have IGRA testing performed, and among IGRA-positive contacts, 18 progressed to TB; among IGRA-negative contacts, 6 progressed to TB. One contact without IGRA testing progressed to TB (**Supplementary Figure 2**). Of the 25 (1.4%) contacts who developed TB, 13 had microbiologically confirmed TB, 2 had histopathological evidence suggestive of TB, and 10 had clinical TB. In addition, 19 (75%) had pulmonary TB and 6 (24%) had extrapulmonary TB.

Clinical characteristics of TB contacts (and their index cases) did not significantly differ between those who did vs. did not progress to TB, except TB progressors were more likely to be IGRA-positive than non-progressors (75% vs. 36.3%). More details are in **Table 1.**

Dimension reduction

For this study, 30 variables were collected at baseline and evaluated: 19 among TB contacts and 11 among TB patients. After empirical review, 22 variables were selected as candidates for the (time-to-event) LASSO regression; in this stage, three variables of contacts were retained: TPT, smoking status, and positive baseline IGRA (**Figure 1**). The inclusion and exclusion of variables in each stage are in **Supplementary Table 1.** For the PCA-LASSO, which used PCA between empirical review and LASSO, two variables were retained: TPT and positive baseline IGRA. More details for the PCA-LASSO are in **Supplementary Table 2.**

Determinants for progression to TB among the entire TB contact population

The final prediction model obtained from LASSO suggested that TPT, smoking status, and baseline IGRA result were predictive of progression to active TB. Internal validation of the model showed an (optimism-corrected) AUC of 0.81 (95%CI: 0.74–0.88) (**Figure 2B**). The results of the post-selection Cox regression demonstrated that contacts that never started TPT were at higher risk of progressing to active TB (aHR=16.55, 95% CI: 2.20-124.43, p-value=0.006), compared to contacts who completed TPT. (**Figure 2A**).

TB progression among contacts who were IGRA-positive at baseline.

Among the 678 contacts who were IGRA-positive at baseline, 18 progressed to TB. The univariable analysis showed that 77.8% of TB progressors never started TPT and only 5.6% completed TPT ($p<0.001$). No other significant differences were found between these groups. (**Supplementary Table 3**). Dimension reduction employing empirical review-LASSO (**Supplementary Table 4**) was performed; of 29 initial variables, only TPT was retained for the final model. The optimism-corrected AUC of this model was 0.73 (95%CI: 0.64–0.79). Post-selection Cox regression showed that never starting TPT was associated with progression to TB (aHR=14.75, 95%CI: 1.95-111.68, p-value=0.009) (**Figure 3**). The same variables were selected using PCA-LASSO (**Supplementary Table 5)**.

Progression to TB among contacts who were IGRA-positive and did not receive TPT or received TPT for <30 days.

Among the 285 contacts who were IGRA-positive at baseline but did not start TPT or received TPT for <30 days, 15 progressed to TB disease. When comparing characteristics of these contacts by TB status, we found that median BMI of TB progressors (24.2 kg/m^2, interquartile range [IQR]: 20.4-28.1) was significantly lower than that of non-progressors (25.3 kg/m^2, IQR: 22.1-36.6) (p-value:0.04) (**Supplementary Table 6**). Empirical review-LASSO dimension reduction retained 2 of 28 potential predictors: BMI of contacts and passive smoking in TB index patients (**Supplementary Table 7**); PCA-LASSO resulted in a similar model (**Supplementary Table 8**). The optimism-corrected AUC of this model was 0.70 (95%CI: 0.59–0.80). In the post-selection Cox model, we found that for each unit increase of BMI, the risk of progressing to TB decreased by 13% (aHR=0.89, 95%CI: 0.80-0.99, p-value=0.027) (**Figure 4**). Schoenfeld residuals are presented in **Supplementary Figure 3.**

We performed additional analyses using restricted cubic splines and segmented logistic regression to identify the critical value of BMI that led to the highest change in the odds of progressing to active TB. The optimal breakpoint was estimated at 24.93 kg/m^2 (95%CI: 20.25- 29.61) **(Figure 5A).** Using this critical point, we found that those contacts with BMI <25 kg/m^2 were more likely to progress to active TB (aHR=4.14, 95%CI: 1.17-14.67, p-value=0.028) **(Figure 5B).**

Moreover, we calculated the probability of progressing to active TB in the group of contacts who were IGRA-positive at baseline and did not receive TPT or received TPT for <30 days: 5.26% (15 progressors/285 contacts). Among IGRA-positive contacts with BMI <25 kg/m^2 who did not receive TPT or received TPT <30 days, the TB risk was 8.4% (12 progressors/143 contacts); in such contacts with BMI ≥25 kg/m^2 the probability was 2.1% (3 progressors/141 contacts). In addition, the breakpoint of 18.5 kg/m^2 was not associated with progression to TB (HR=1.23, 95%CI: 0.39-3.92), using 25 kg/m^2 as the reference level.

We also evaluated the effect of BMI on TB progression via Kaplan-Meier curves (**Supplementary Figure 4**) and via Cox models with an added interaction term between BMI and TPT (**Supplementary Table 9**), for all three study groups. All additional analyses suggested a protective effect of BMI on TB progression.

We used five-fold cross-validation to build all three prediction models. Sensitivity analyses using 10-fold and leave-one-out-cross validation were carried out, resulting in similar results (**Supplementary Figure 5 and 6**, respectively). Likewise, similar predictors were selected after using empirical review, PCA and LASSO **(Supplementary Table 10, 11 and 12)**. Sensitivity analyses using complete case analyses led to similar results for each group of interest (**Supplementary Table 13, 14 and 15)**.

DISCUSSION

We found that close TB contacts who were IGRA-positive at baseline but did not receive TPT and had lower BMI (i.e., <25 kg/m^2) were at increased risk of progression to TB. Of note, 12 of 15 (80%) of those who progressed to TB were identified by this algorithm. Previous studies have reported being underweight as a risk factor for incident active TB [8] while overweight and normal weight individuals had a reduced risk of TB disease [28]. In a recent post-hoc analysis of a clinical trial among people with HIV treated with TPT (not close contacts; 38% were IGRA-positive), BMI <25 kg/m^2 was also associated with increased TB risk [29]. Our analyses of IGRA-positive close TB contacts who did not receive TPT (none of whom had HIV) is particularly relevant. The BMI value identified here coincides with the upper value for the normal BMI classification (18-25 kg/m^2).

Previous studies have shown that overweight and obese individuals (though not specifically close contacts of people with TB) have a lower risk of TB, compared to those with normal BMI and low BMI [7, 8]. Although the impact of immunological and biological mechanisms of BMI on TB progression is complex, nutritional differences between overweight or obese individuals and those with normal or low weight may be one of the possible factors explaining the protective effect against TB progression [7]. Vitamin A [30] and E [31] deficiency have both been associated with an increased risk of progression to TB among household TB contacts.

Low BMI can influence innate and adaptive immune responses through various mechanisms [32]. There is evidence of a correlation between low BMI in people with pulmonary TB and decreased levels of IFNγ, TNFα, IL-2, IL-17, IL-6, and IL-12, affecting protective immunity [32]. Reduced chemokine levels have also been observed in individuals with Mtb infection [32]. Moreover, Mtb can use adipose tissue to survive without being recognized by the host immune system [33], which would explain why overweight or obesity are protective factors for the development of TB [33]. Corroborating this, a study of cost-effectiveness models for reducing TB incidence and mortality in India showed that a robust nutritional intervention could prevent 78% of incident TB diagnoses and 88% of deaths among malnourished close contacts [34]. Furthermore, a clinical trial also in India demonstrated a reduced incidence of pulmonary TB (48%) among household contacts of TB patients who received nutritional support [35].

Using characteristics of close TB contacts and their index cases, we performed a prediction model for progression to TB at the time of evaluation of the close contact. Among close contacts with Mtb infection (i.e., IGRA-positive), BMI was the only variable that predicted progression to TB. Among IGRA-positive close contacts who did not receive TPT or received TPT <30 days, BMI <25 kg/m^2 was associated with an approximately 4-fold increased odds of progressing to TB. In addition, the number of contacts needing treatment decreased from 285 to 143, a 50% reduction.

In the prediction model among all close contacts, as well as among IGRA-positive close contacts, no TPT start was the strongest risk factor for progression to TB. TPT is highly effective in

preventing progression to TB disease [36, 37]. Two studies, one in Pakistan [37] and another using a decision analysis model [36] reported that contact investigation programs related to the initiation of TPT prevent progression to TB disease compared to no intervention, varying the number of cases averted according to the type of contact (age and type of exposure), TPT scheme and detection algorithm [36, 37].

In 2024, the Word Health Organization emphasized the recommendation to provide TPT to all household contacts of all ages [1], with many countries expanding the provision of TPT from children <5 years to children <15 years regardless of IGRA result. Despite these recommendations, in 2023, only 7% of the TPT household contacts aged <5 years and 22% of contacts ≥5 years received TPT [1]. Thus, TPT implementation remains deficient among contacts of all ages.

In the full cohort (regardless of IGRA and TPT status) TB incidence among contacts was 1.4%. A recent systematic review [38] found an incidence of active TB among household contacts of 2% in the first year and 0.75% in the second year following exposure; incidence declined in subsequent years, regardless of IGRA and TPT status. Among close contacts in this study who were IGRA-positive and did not receive TPT, 5% progressed to TB. This is consistent with several previous studies [39, 40].

This study had several limitations. First, we did not obtain information on the type of contacts (e.g., household, coworkers), the spatial proximity of contacts to index cases, nor access to important social resources such as Bolsa Familia payments. Second, it was not possible to determine whether the contact became infected after exposure or if there was pre-existing Mtb infection. Third, the number of TB progressors was low, limiting our ability to fit richer and more complex prediction models; however, the proportion of progressors was consistent with the existing literature (reinforcing external validity): 1.4% of all contacts and 5% of IGRA-positive contacts who did not receive preventive therapy [39, 40]. In addition, the dimension reduction techniques utilized, such as LASSO, have been shown to perform well even with few events [22]. Fourth, the proportion of study participants with BMI <18.5 kg/m^2 was relatively low (330/1845; 17.9%), and cultures were not obtained in individuals without symptoms. However, none of the TB progressors reported weight loss or had an abnormal baseline chest X-ray (despite all contacts having a baseline chest X-ray available), making it less likely that they had undiagnosed TB, including asymptomatic TB. Finally, external validation of the main prediction model could not be performed. Confirmation of this BMI-based approach, and the BMI cutoff value of 25 kg/m^2, in a separate cohort would provide important evidence before using this predictor in clinical practice. However, the dimension reduction approach at each step, from the selection of variables to the final mixed-effects Cox proportional hazard analysis, was exhaustive, and we achieved robust internal validation in a large, diverse cohort.

In conclusion, BMI<25 kg/m^2 identified IGRA-positive close TB contacts who were at high risk of progressing to TB disease. An emphasis on TPT in this high-risk group could facilitate TB prevention efforts.

NOTES

Acknowledgements: The authors thank the study participants. We also thank the clinical and laboratory teams of RePORT Brazil.

Disclaimer: The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data availability: The data used in the manuscript is available upon request to the corresponding author.

Contributions: Conceptualization, M.C.F., M.C.S., V.C.R., B.B.A and T.R.S..; Data curation, M.B.A., G.A., M.C.F., C.S. and T.R.S.; Investigation, M.B.A., G.A, M.C.F., C.S., and T.R.S. ;Formal analysis, M.B.A., P.F.R. and G.A. ; Funding acquisition, A.L.K., M.C.S., M.C.F., V.C.R., B.B.A and T.R.S.; Methodology, M.B.A., G.A., and T.R.S.; Project administration, M.C.F., C.S and T.R.S.; Resources, M.B.A., G.A., M.C.F., C.S. and T.R.S.; Software, M.B.A., G.A., and T.R.S.; Supervision, P.F.R. and T.R.S.; Writing—original draft, M.B.A., G.A., P.F.R. and T.R.S.; Writing review and editing, all authors. All authors have read and agreed to the submitted version of the manuscript.

Funding: This work was supported by the Project Number R01-AI147765 from NIH to M. B. A. The work from B.B.A., M.C.S., A.L.K., M.C.F., V.C.R., G.A. and T.R.S was supported by intramural research program from FIOCRUZ, by the National Institutes of Health (U01-AI069923) and by the Departamento de Ciência e Tecnologia (DECIT) - Secretaria de Ciência e Tecnologia (SCTIE) – Ministério da Saúde (MS), Brazil (25029.000507/2013–07). B.B.A., M.C.S., A.L.K. and V.C.R. are senior scientists from the Conselho Nacional de Desenvolvimento Científico e Tecnológico, and M.B.A is Researcher Level I from The Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (CONCYTEC).

Potential conflicts of interest: All authors no reported conflicts of interest.

References

1. World Health Organization. Global Tuberculosis Report 2024. Genova: World Health Organization, **2024**.
2. Getahun H, Matteelli A, Chaisson RE, Raviglione M. Latent Mycobacterium tuberculosis infection. N Engl J Med **2015**; 372(22): 2127-35.
3. Houben RM, Dodd PJ. The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. PLoS Med **2016**; 13(10): e1002152.
4. Reichler MR, Khan A, Sterling TR, et al. Risk and Timing of Tuberculosis Among Close Contacts of Persons with Infectious Tuberculosis. J Infect Dis **2018**; 218(6): 1000-8.
5. Trauer JM, Moyo N, Tay EL, et al. Risk of Active Tuberculosis in the Five Years Following Infection . . . 15%? Chest **2016**; 149(2): 516-25.

6. Edwards LB, Livesay VT, Acquaviva FA, Palmer CE. Height, weight, tuberculous infection, and tuberculous disease. Arch Environ Health **1971**; 22(1): 106-12.
7. Lonnroth K, Williams BG, Cegielski P, Dye C. A consistent log-linear relationship between tuberculosis incidence and body mass index. Int J Epidemiol **2010**; 39(1): 149-55.
8. Saunders M, Cegielski J, Clark R, Houben R, McQuaid C. Tuberculosis Incidence and Body Mass Index – An Updated Systematic Literature Review and Global Analysis. Available at: https://ssrn.com/abstract=5032687.
9. Guwatudde D, Nakakeeto M, Jones-Lopez EC, et al. Tuberculosis in household contacts of infectious cases in Kampala, Uganda. Am J Epidemiol **2003**; 158(9): 887-98.
10. Moran-Mendoza O, Marion SA, Elwood K, Patrick D, FitzGerald JM. Risk factors for developing tuberculosis: a 12-year follow-up of contacts of tuberculosis cases. Int J Tuberc Lung Dis **2010**; 14(9): 1112-9.
11. Yoshiyama T, Harada N, Higuchi K, Saitou M, Kato S. Use of the QuantiFERON(R)-TB Gold in Tube test for screening TB contacts and predictive value for active TB. Infect Dis (Lond) **2015**; 47(8): 542-9.
12. Reichler MR, Khan A, Sterling TR, et al. Risk Factors for Tuberculosis and Effect of Preventive Therapy Among Close Contacts of Persons With Infectious Tuberculosis. Clin Infect Dis **2020**; 70(8): 1562-72.
13. Li R, Nordio F, Huang CC, et al. Two Clinical Prediction Tools to Improve Tuberculosis Contact Investigation. Clin Infect Dis **2020**; 71(8): e338-e50.
14. Davies LRL, Wang C, Steigler P, et al. Age and sex influence antibody profiles associated with tuberculosis progression. Nat Microbiol **2024**; 9(6): 1513-25.
15. Park SY, Han S, Kim YM, et al. Risk of active tuberculosis development in contacts exposed to infectious tuberculosis in congregate settings in Korea. Sci Rep **2020**; 10(1): 1306.
16. Souza AB, Arriaga MB, Amorim G, et al. Determinants of losses in the latent tuberculosis infection cascade of care in Brazil. BMJ Glob Health **2021**; 6(9): e005969.
17. Mendelsohn SC, Andrade BB, Mbandi SK, et al. Transcriptomic signatures of progression to TB disease among close contacts in Brazil. J Infect Dis **2024**: e1355-e65.
18. Arriaga MB, Amorim G, Queiroz ATL, et al. Novel stepwise approach to assess representativeness of a large multicenter observational cohort of tuberculosis patients: The example of RePORT Brazil. Int J Infect Dis **2021**; 103: 110-8.
19. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform **2009**; 42(2): 377-81.
20. Ministério da Saúde do Brasil, Secretaria de Vigilância em Saúde. Manual de Recomendações para o Controle da Tuberculose no Brasil. Brazil, **2019**:288.
21. Pavlou M, Ambler G, Seaman S, De Iorio M, Omar RZ. Review and evaluation of penalised regression methods for risk prediction in low-dimensional data with few events. Stat Med **2016**; 35(7): 1159-77.
22. Rafati S, Baneshi MR, Hassani L, Bahrampour A. Comparison of Penalized Cox Regression Methods in Low-Dimensional Data with Few-Events: An Application to Dialysis Patients' Data. J Res Health Sci **2019**; 19(3): e00452.

23. Rickles M, Rebeiro PF, Sizemore L, et al. Tennessee's In-state Vulnerability Assessment for a "Rapid Dissemination of Human Immunodeficiency Virus or Hepatitis C Virus Infection" Event Utilizing Data About the Opioid Epidemic. Clin Infect Dis **2018**; 66(11): 1722-32.
24. Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med **1997**; 16(4): 385-95.
25. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med **1996**; 15(4): 361-87.
26. Lee J, Sun D, Sun Y, Taylor J. Exact post-selection inference, with application to the lasso. The Annals of Statistics **2016**; 44: 907-27.
27. Van Buuren S. Flexible imputation of missing data: CRC press, **2018**.
28. Hanrahan CF, Golub JE, Mohapi L, et al. Body mass index and risk of tuberculosis and death. AIDS **2010**; 24(10): 1501-8.
29. Nguenha D, Ndebele F, Saavedra B, et al. BMI as a predictor of progression from TB infection to active TB in PLHIV. Int J Tuberc Lung Dis **2025**; 29(2): 54-9.
30. Aibana O, Franke MF, Huang CC, et al. Impact of Vitamin A and Carotenoids on the Risk of Tuberculosis Progression. Clin Infect Dis **2017**; 65(6): 900-9.
31. Aibana O, Franke MF, Huang CC, et al. Vitamin E Status Is Inversely Associated with Risk of Incident Tuberculosis Disease among Household Contacts. J Nutr **2018**; 148(1): 56-62.
32. Kumar NP, Nancy AP, Moideen K, et al. Low body mass index is associated with diminished plasma cytokines and chemokines in both active and latent tuberculosis. Front Nutr **2023**; 10: 1194682.
33. Neyrolles O, Hernandez-Pando R, Pietri-Rouxel F, et al. Is adipose tissue a place for Mycobacterium tuberculosis persistence? PLoS One **2006**; 1(1): e43.
34. Sinha P, Lakshminarayanan SL, Cintron C, et al. Nutritional Supplementation Would Be Cost-Effective for Reducing Tuberculosis Incidence and Mortality in India: The Ration Optimization to Impede Tuberculosis (ROTI-TB) Model. Clin Infect Dis **2022**; 75(4): 577-85.
35. Bhargava A, Bhargava M, Meher A, et al. Nutritional supplementation to prevent tuberculosis incidence in household contacts of patients with pulmonary tuberculosis in India (RATIONS): a field-based, open-label, cluster-randomised, controlled trial. Lancet **2023**; 402(10402): 627-40.
36. Doan TN, Fox GJ, Meehan MT, et al. Cost-effectiveness of 3 months of weekly rifapentine and isoniazid compared with other standard treatment regimens for latent tuberculosis infection: a decision analysis study. J Antimicrob Chemother **2019**; 74(1): 218-27.
37. Kendall EA, Hussain H, Kunkel A, et al. Isoniazid or rifampicin preventive therapy with and without screening for subclinical TB: a modeling analysis. BMC Med **2021**; 19(1): 315.
38. Velen K, Shingde RV, Ho J, Fox GJ. The effectiveness of contact investigation among contacts of tuberculosis patients: a systematic review and meta-analysis. Eur Respir J **2021**; 58(6).
39. Sutherland I. Recent studies in the epidemiology of tuberculosis, based on the risk of being infected with tubercle bacilli. Adv Tuberc Res **1976**; 19: 1-63.
40. Kritski AL, Marques MJ, Rabahi MF, et al. Transmission of tuberculosis to close contacts of patients with multidrug-resistant tuberculosis. Am J Respir Crit Care Med **1996**; 153(1): 331-5.

TABLES

Table 1. Characteristics of TB contacts and TB patients by progression to TB status.

Characteristics	n/N	No progression to TB (n=1821)	Progression to TB (n=25)	p-value
TB contacts				
Age – median (IQR)	1846/1846	31.8 (16.1-64.9)	31.9 (20.2-71.1)	0.33
Female– no. (%)	1846/1846	1075 (59)	16 (64)	0.69
Race/Ethnicity – no. (%)	1844/1846			0.24
White		356 (19.6)	6 (24)	
Black		363 (20)	7 (28)	
Asian		5 (0.3)	0 (0)	
Pardo		1081 (59.4)	12 (48)	
Indigenous		14 (0.8)	0 (0)	
BMI – median (IQR)	1845/1846	24.3 (20.2-36)	24.2 (19.6-32.2)	0.61
BMI- categories– no. (%)	1845/1846			0.26
< 18.5		326 (17.9)	4 (16)	
18.5 - 25		686 (37.7)	13 (52)	
> 25		807 (44.4)	8 (32)	
Literate– no. (%)	1843/1846	1635 (89.9)	23 (92)	0.73
Income – no. (%)	1784/1846			0.18
Without income		373 (21.2)	6 (26.1)	
Equal or less than a minimum wage		717 (40.7)	12 (52.2)	
More than a minimum wage		671 (38.1)	5 (21.7)	
BCG scar – no. (%)	1845/1846	1631 (89.6)	21 (84)	0.32
HIV infection – no. (%)	1844/1846	48 (2.6)	0 (0)	>0.99
ART (if HIV+) – no. (%)	48/48	38 (79.2)	0 (0)	0.22
Smoking – no. (%)	1845/1846			0.20
Current		188 (10.3)	1 (4)	
Former		296 (16.3)	3 (12)	
Never		1336 (73.4)	21 (84)	
Passive smoking – no. (%)	1838/1846	585 (32.3)	10 (40)	0.40
Alcohol consumption – no. (%)	1845/1846			0.12
Current		624 (34.3)	6 (24)	
Former		350 (19.2)	3 (12)	
Never		846 (46.5)	16 (64)	
CAGE assessment – no. (%)	982/1846	171 (17.6)	3 (33.3)	0.20
Illicit drug use – no. (%)	1845/1846			0.10
Current		43 (2.4)	0 (0)	

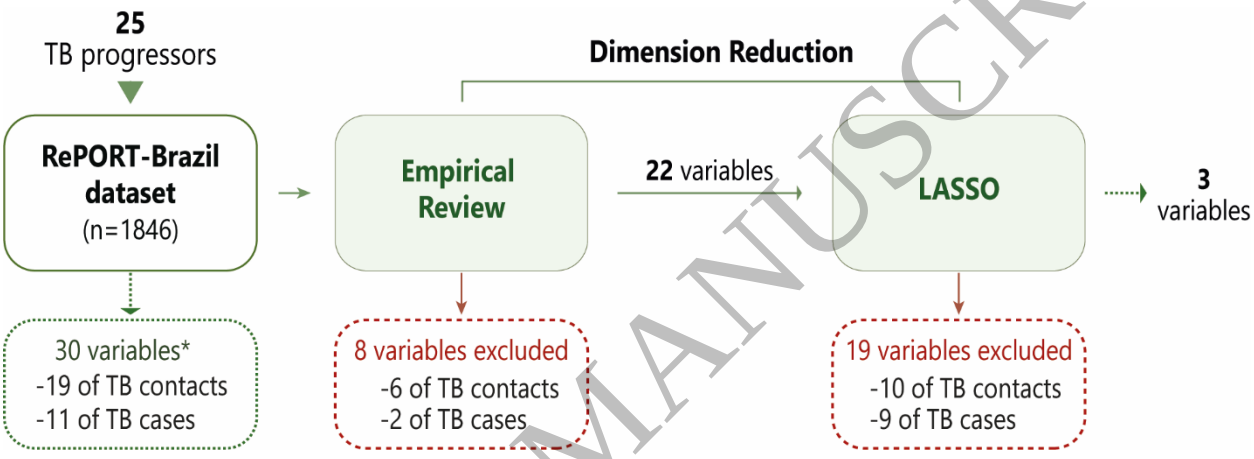
Former		148 (8.1)	0 (0)	
Never		1629 (89.5)	25 (100)	
Abnormal chest X-ray – no. (%)		8 (0.4)	0 (0)	NA
Diabetes – no. (%)	1821/1846	89 (5)	1 (4)	>0.99
Hypertension – no. (%)	1821/1846	219 (12.2)	3 (12)	>0.99
Time index contact (weeks) – median (IQR)	1821/1846	37 (13-255)	20 (5-144)	0.06
TPT status – no. (%)	1846/1846			0.44
Not recommended		983 (54)	4 (16)	
Recommended, never started		352 (19.3)	16 (64)	
Recommended, incomplete TPT		168 (9.2)	4 (16)	
Recommended, complete TPT		318 (17.5)	1 (4)	
IGRA result at baseline– no. (%)	1841/1846			**<0.01**
Negative		1126 (62)	6 (25)	
Positive		660 (36.3)	18 (75)	
Indeterminate		31 (1.7)	0 (0)	
TB index patients				
Age – median (IQR)	1846/1846	34 (24-64)	38 (26-64)	0.71
Female– no. (%)	1842/1846	1126 (62)	17 (68)	0.68
Multidrug resistance – no. (%)	1822/1846	29 (1.6)	0 (0)	>0.99
HbA1c (%)-median (IQR)	1839/1846	5.9 (5.5-12.1)	5.8 (5.3-13.8)	0.71
Cough at baseline – no. (%)	1842/1846	1682 (92.6)	25 (100)	0.25
Abnormal chest X-ray– no. (%)	1647/1846	833 (51.4)	15 (60)	0.43
Smoking – no. (%)	1842/1846			0.77
Current		325 (17.9)	4 (16)	
Former		598 (32.9)	8 (32)	
Never		894 (49.2)	13 (52)	
Passive smoking – no. (%)	1830/1846	1190 (65.9)	17 (68)	>0.99
Alcohol consumption – no. (%)	1842/1846			0.66
Current		680 (37.4)	10 (40)	
Former		793 (43.6)	8 (32)	
Never		344 (18.9)	7 (28)	
Illicit drug use – no. (%)	1842/1846			0.29
Current		200 (11)	2 (8)	
Former		428 (23.6)	4 (16)	
Never		1189 (65.4)	19 (76)	
Positive sputum smear – no. (%)	1820/1846	1373 (76.5)	20 (80)	0.82

Table note. Data represent no. (%), except for age, BMI and HbA1c, which is presented as median and interquartile range (IQR). Continuous variables were compared using the Mann Whitney U test and categorical variables were using the Pearson's chi-square test. Bold values represent statistically significant.

Abbreviations: TB: tuberculosis, BMI: Body Mass Index, ART: antiretroviral therapy, IGRA; Interferon-Gamma Release Assay, TPT: Tuberculosis preventive treatment. NA: not applicable. CAGE assessment: Cut down, Annoyed, Guilty, and Eye-opener, test for alcohol use disorder.NA: not applicable

FIGURES AND LEGENDS

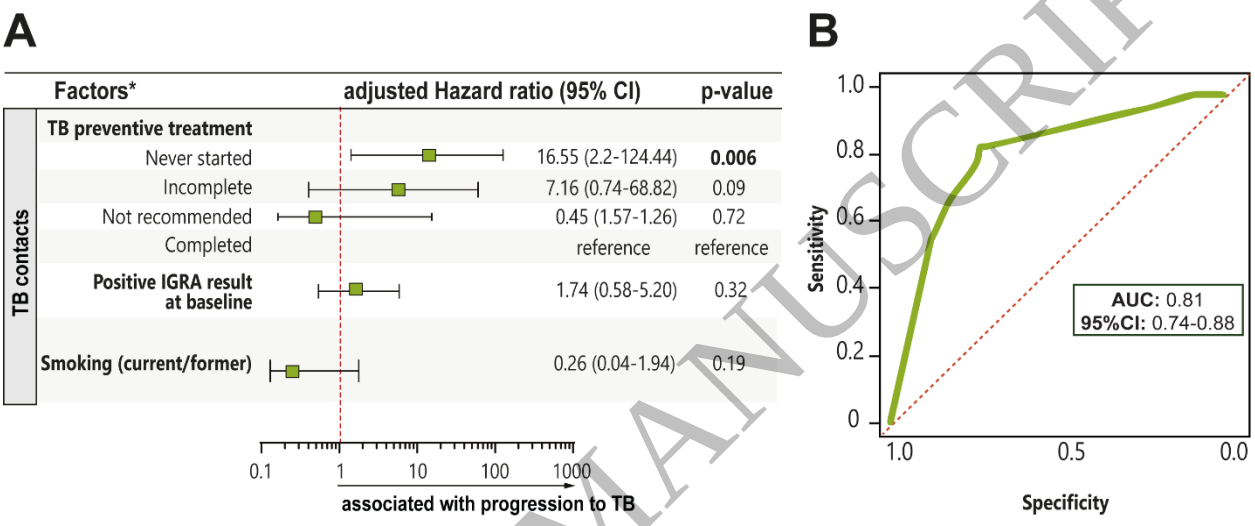
Figure 1. Dimension reduction stages. This scheme shows the number of variables initiated, excluded, and retained at empirical review and LASSO regression stages. More details are in **Supplementary Table 1.**



*Data collected at baseline.

Abbreviations: TB: Tuberculosis, CAGE assessment+: Cut down, Annoyed, Guilty, and Eye-opener, test for alcohol use disorder.

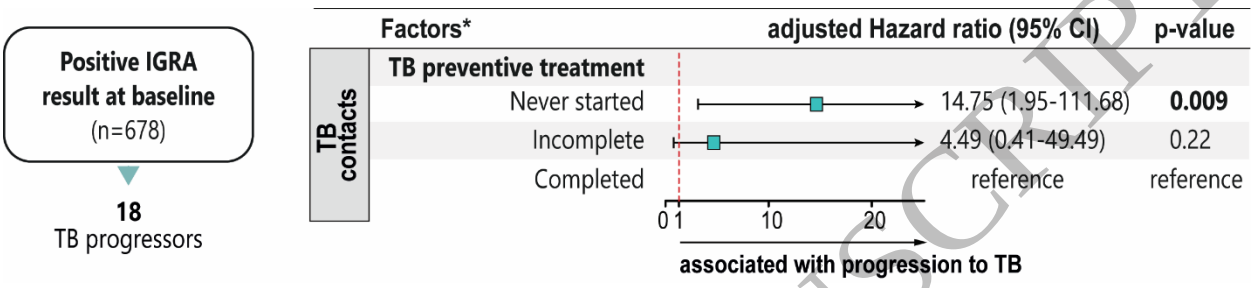
Figure 2. Determinants for progression to TB among the entire TB contact population. (A) Mixed-effects Cox proportional hazard model was performed to test association between characteristics of TB contacts or pulmonary TB patients and progression to TB among all close contacts enrolled in the study. Dimension reduction was used to select characteristics of TB contacts and TB patients included in this model. **(B)** Receiver operator characteristics (ROC) curve for model development using 200 bootstraps. More details are in **Table 1** and **Supplementary Table 1.**



*Data collected at baseline

Abbreviations: IGRA: Interferon-Gamma Release Assay, TB: Tuberculosis, TPT: Tuberculosis preventive treatment, Time (index -contact): time difference between visit 1 of the TB index patient and visit 1 of the contact, AUC: area under the curve, CI: confidence interval

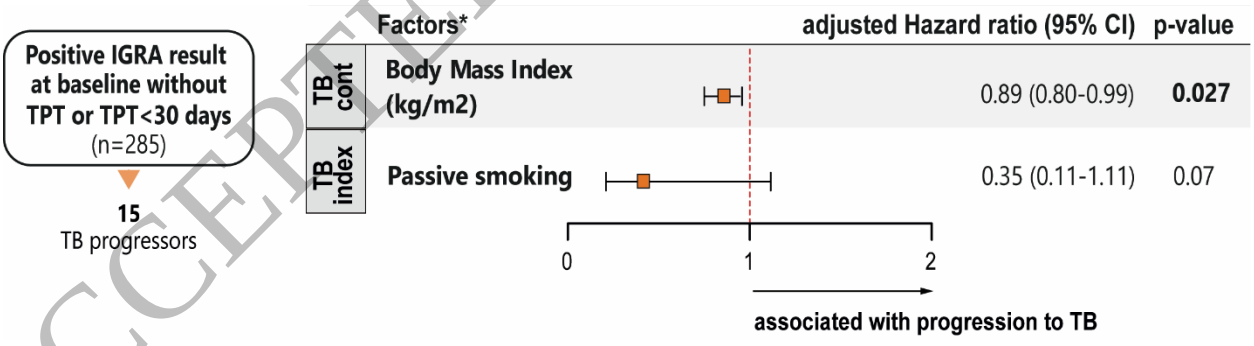
Figure 3. Determinants for progression to TB among contacts who were IGRA-positive at baseline. Mixed-effects Cox proportional hazard model was performed to test the association between characteristics of TB contacts or pulmonary TB patients and progression to TB among contacts with positive IGRA result at baseline. Dimension reduction was used to select characteristics of TB contacts and TB patients included in this model. More details are in **Supplementary Table 3 and 4.**



*Data collected at baseline

Abbreviations: IGRA: Interferon-Gamma Release Assay, TB: Tuberculosis, TPT: Tuberculosis preventive treatment, CI: confidence interval

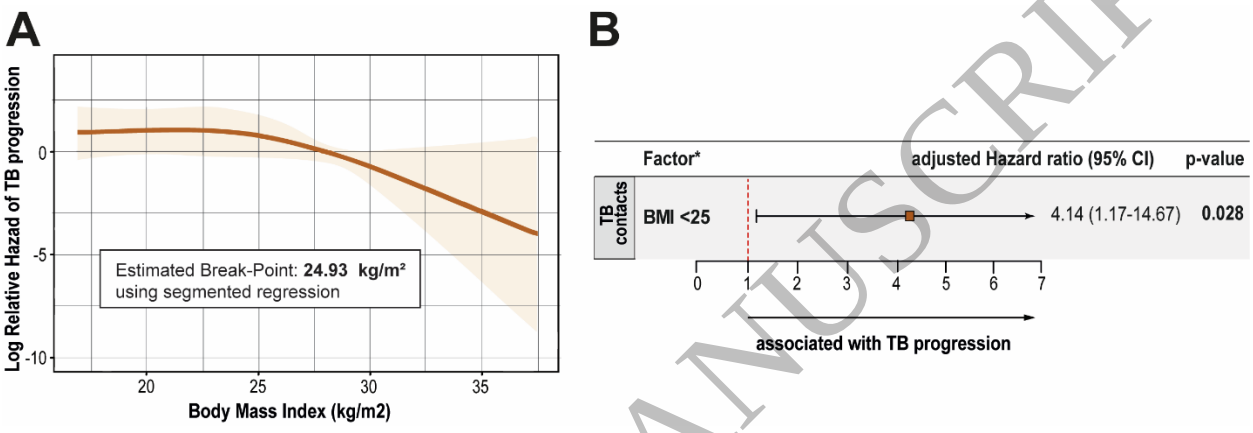
Figure 4. Determinants for progression to TB among contacts who were IGRA-positive and did not receive TPT or received TPT for <30 days. Mixed-effects Cox proportional hazard model was performed to test associations between characteristics of TB contacts or pulmonary TB index patients and progression to TB among contacts with a positive IGRA result at baseline who did not start TPT or received <30 days. Dimension reduction was used to select the variables included in this model. More details are in Supplementary Table 6 and 7.



*Data collected at baseline

Abbreviations: IGRA: Interferon-Gamma Release Assay, TB: Tuberculosis, CI: confidence interval

Figure 5. Body Mass Index as a determinant for progression to TB among contacts who were IGRA-positive and did not receive TPT or received TPT for <30 days. (A) Regression Model with Segmented Relationship to find the optimal break-point using BMI values of contacts with IGRA-positive result at baseline and who did not start treatment or received <30 days. (B) Logistic regression using the optimal break-point BMI among contacts with positive IGRA result at baseline and who did not start TPT or received <30 days. BMI ≥ 25 kg/m^2 was used as the reference.



*Data collected at baseline

Abbreviations: IGRA: Interferon-Gamma Release Assay, TB: Tuberculosis, CI: confidence interval