Tuberculosis symptom screening for children and adolescents living with HIV in six high HIV/TB burden countries in Africa

Bryan Vonasek\textsuperscript{a,b}, Alexander Kay\textsuperscript{a,c,d}, Tara Devezina\textsuperscript{a,c}, Jason M. Bacha\textsuperscript{a,e,f}, Peter Kazembe\textsuperscript{a,b,f}, Dilsher Dhillon\textsuperscript{a,c}, Sandile Dlamini\textsuperscript{d,f}, Heather Haq\textsuperscript{a}, Lineo Thahane\textsuperscript{a,f,g}, Katie Simon\textsuperscript{a,b,f}, Mogomotsi Matshaba\textsuperscript{a,f,h}, Jill Sanders\textsuperscript{f,g}, Mercy Minde\textsuperscript{i}, Sebastian Wanless\textsuperscript{f}, Phoebe Nyasulu\textsuperscript{b} and Anna Mandalakas\textsuperscript{a,c}

Objectives: The WHO recommends that children and adolescents living with HIV (CALHIV) complete TB symptom screening at every clinical encounter but evidence supporting this recommendation is limited. We evaluated the performance of the recommended TB symptom screening in six high-burden TB/HIV countries.

Design: Retrospective longitudinal cohort.

Methods: We extracted data from electronic medical records of CALHIV receiving care from clinics in Botswana, Eswatini, Lesotho, Malawi, Tanzania, and Uganda from January 2014 to June 2017. We defined incident TB cases as those prescribed TB treatment within 30 days of TB diagnosis. We analyzed the most recent symptom screen preceding a TB diagnosis. In accordance with WHO guidelines, positive screens were defined as current fever, cough, poor weight gain, or recent TB contact. Odds of TB disease was modeled by screen result and age at which screening was conducted.

Results: Twenty thousand seven hundred and six patients collectively had 316,740 clinic visits, of which 240,161 (75.8\%) had documented TB symptom screens. There were 35,701 (14.9\%) positive TB symptom screens, and 1212 incident TB diagnoses. Sensitivity and specificity of the TB symptom screen to diagnose TB were 61.2\% (95\% CI 58.4–64.0) and 88.8\% (95\% CI 88.7–88.9), respectively. Log odds of documented TB for positive or negative screens was statistically different only for screens conducted at ages 7–17.

Conclusion: Although specificity was high, the sensitivity of the TB symptom screen to detect TB in CALHIV was low. More accurate screening approaches are needed to optimally identify TB disease in CALHIV.

Keywords: adolescents, children, HIV, intensive case finding, tuberculosis
Introduction

Early detection of tuberculosis (TB) infection and disease significantly reduces mortality in children and adolescents living with HIV (CALHIV) [1,2]. To improve early detection, the WHO recommends using symptom screening as part of intensive case finding (ICF) efforts in CALHIV presenting to health facilities. For children living with HIV, this was classified as a ‘strong recommendation’, but with ‘low quality evidence’ by the WHO [3]. For adolescents living with HIV, this was classified as a ‘strong recommendation’ with ‘moderate quality evidence’ [3]. A study of HIV-positive adults in Uganda that compared outcomes before and after implementation of ICF symptom screening demonstrated that this strategy was associated with only limited improvements in detection of TB and led to significantly higher estimated costs primarily because of false-positive screens [4]. Yoon et al. [5] recently highlighted the insufficient sensitivity of relying on assessment of symptoms for TB screening and emphasized the need to incorporate testing into screening strategies as a leap forward in ending the global TB epidemic. With children typically having even less specific symptoms for TB, there is concern that the ICF symptom screening may have less accuracy in this age group [3]. However, we are only aware of one study that has directly assessed the accuracy of the ICF symptom screening strategy in the pediatric HIV population [6], and additional insight into the predictive utility of this screening strategy will help clinicians interpret screening results and inform future recommendations for screening in this high-risk population.

To expand the evidence base for or against the WHO-recommended ICF symptom screening strategy for CALHIV, the aim of this study was to assess the performance of the screening strategy at seven sites across six high TB/HIV-burden countries in eastern and southern Africa.

Methods

This was a retrospective longitudinal cohort analysis of data available from the electronic medical records (EMR) of seven Baylor International Pediatric AIDS Initiative (BIPAI) sites in Africa (Gaborone, Botswana; Mbabane, Eswatini; Maseru, Lesotho; Lilongwe, Malawi; Mbeya, Tanzania; Mwanza, Tanzania; and Kampala, Uganda). The main priority of these clinics is to provide free, comprehensive care to CALHIV, including diagnosis and management of HIV-associated TB. Since 2014, these sites have routinely implemented TB symptom screening based upon the WHO ICF recommendations. EMR data from January 2014 to June 2017 was abstracted and analyzed for all sites except for Lilongwe. In Lilongwe, data was extracted from April 2016 to June 2017 as documentation of TB symptom screening started later at this site.

In accordance with WHO ICF guidelines [3], any current fever, any cough, poor weight gain, or recent TB contact defined a positive screen. At these sites, the large majority of TB symptom screens are conducted by lay screening officers prior to the visit with a clinician. The clinician then completes the TB assessment with the result of the screen from that visit and any previous visits at that site available in the EMR. TB assessments and clinical visits are conducted by physicians, nonphysician clinicians, and nurses at these sites. All of these clinicians have been trained and oriented to make TB disease classification consistent with the WHO Definitions and Reporting Framework for Tuberculosis [7]. The large majority of TB disease diagnoses for a particular visit are made by, or made after discussion with, an experienced physician or nonphysician clinician. Despite diagnostic evaluations supported by appropriate child TB specimen collection, Gene Xpert MTB/RIF testing, and radiographic imaging, most child TB cases at these sites are clinically diagnosed rather than bacteriologically confirmed, which is typical in all settings. Tests of TB infection, such as tuberculin skin tests or interferon-gamma release assays, are not routinely conducted at these sites.

Patient results were included in this analysis if the patient was under 20 years of age and enrolled in care at the site with documented HIV infection. The TB symptom screen (‘positive’, ‘negative’, or missing) and TB assessment (‘TB disease’, ‘presumptive TB’, ‘no TB’, or missing) from each clinic visit for patients was assessed. Visits with no documented TB symptom screen or TB assessment were excluded. ‘Presumptive TB’ is not a definitive diagnosis and captures patients completing ongoing consideration for TB disease in whom TB is either ruled out or diagnosed definitively at future visits. Hence, visits with ‘presumptive TB’ were excluded from analysis. Incident cases of TB were defined as the visit with the first documented assessment of ‘TB disease’. Only incident cases with documented initiation of antituberculosis therapy within 30 days were included. This was done to avoid inclusion of erroneously documented ‘TB disease’. Prior to the study period, documentation of TB diagnosis and management was not standardized across all sites. To further ensure accurate capture of incident TB cases, data from patients with TB documented before the study period were excluded. TB symptom screens from the visit defining incident TB disease cases were included for calculation of screening test characteristics. If no TB symptom screen was available from the day of an incident case, the most recent screen within 30 days prior was included.

The data were summarized by frequency and percent or mean and interquartile range (IQR). Screening test
characteristics were reported with 95% confidence intervals (CI). A sensitivity analysis assessed the effect on screening performance when including patients with previously documented TB disease or no documented initiation of antituberculosis therapy within 30 days of TB diagnosis. To assess the relationship between the predictive utility of TB symptoms and screening age, we estimated the relationship between the log odds of documented TB diagnosis with age stratified by screening results. Age was allowed to vary nonlinearly by using a penalized cubic spline function using the library ‘gamm4’ in R [8] and placing knots at evenly separated distances across age. The weights were estimated using a mixed model representation employing restricted maximum likelihood estimation across ages. The yield of the screen was defined as a percentage of the number of incident TB disease cases per number of documented screens [9]. Ethical approval was obtained from all necessary ethical bodies in each country including the Baylor College of Medicine Children’s Foundation or Trust, the national ethics committees in each country, and the Baylor College of Medicine Institutional Review Board. Informed consent was not obtained for retrospective review of medical records, and data was deidentified upon extraction from the EMR.

**Results**

The 20,706 CALHIV considered in the analysis collectively completed 316,740 clinic visits that included 240,161 documented TB symptom screens of which 14.9% (35,701 of 240,161) were positive (Fig. 1). Symptom screening was performed in over 90% of clinic visits at all sites with the exception of Botswana, where EMR-based documentation of TB symptom screening was not uniformly adopted until 2015 and screening occurred in as low as 35% of visits in 2014. Fifty percent of screens were completed on female patients. The median

![Diagram](image_url)
The age of patients at the time of screening was 11.2 years (IQR 6.9–15.0 years). Among all patients, 7.7% (1592 of 20,706) had documented TB disease, of which 76.1% (1212 of 1592) fit inclusion criteria for incident TB disease in the primary analysis. TB disease was microbiologically confirmed in 32% of TB cases in this cohort. The yield of the TB symptom screen was 0.6% overall, ranging from 0.3% in Botswana to 1.4% in Lesotho.

We examined several characteristics of the screening test (Table 1). Secondary analysis, including patients with previously documented TB disease or no documented initiation of antituberculosis therapy within 30 days of TB diagnosis, produced estimates that differ compared with the primary analysis. Notably, the estimate of sensitivity was lower and the estimate of positive-predictive value was higher in the secondary analysis. However, these differences were unlikely to be clinically significant.

The relationship between TB symptom screen result and log odds of TB was largely constant across ages at which screening was conducted (Fig. 2). Estimated log odds of TB for negative and positive screens were statistically different only for screens conducted at ages 7–17. Log odds of TB for negative and positive screens was not statistically different for those over 17 years and under 7 years of age, with differentiation especially lacking in the first few years of life.

**Discussion**

Optimizing the WHO ICF strategy is a critical approach for preventing morbidity and mortality in CALHIV given the burden of TB in this population, especially in high-TB/HIV burden countries [2,10–14]. During this 36-month period, routine TB symptom screening was documented at over three-quarters of clinical encounters in this cohort of over 20,000 CALHIV from six high-TB/HIV burden countries in eastern and southern Africa. Accurate understanding of a screening strategy’s performance allows clinicians to better apply the results in context. In CALHIV in our study settings, positive TB symptom screens have low PPV of 3.2% but the positive

---

**Table 1. Tuberculosis symptom screen performance measures.**

| Measure                     | Primary analysis (95% CI) | Secondary analysis (95% CI) |
|-----------------------------|---------------------------|-----------------------------|
| Sensitivity                 | 61.2% (58.4–64.0)         | 54.7% (52.2–57.2)           |
| Specificity                 | 88.8% (88.7–88.9)         | 88.8% (88.6–88.9)           |
| Positive-predictive value   | 3.2% (3.0–3.4)            | 3.7% (3.3–4.0)              |
| Negative-predictive value   | 99.7% (99.7–99.7)         | 99.6% (99.6–99.6)           |
| Positive likelihood ratio   | 5.5 (5.2–5.7)             | 4.87 (4.65–5.10)            |
| Negative likelihood ratio   | 0.44 (0.41–0.47)          | 0.51 (0.48–0.54)            |

Primary analysis was performed with exclusion of patients with previously documented TB disease or no documented initiation of antituberculosis therapy within 30 days of TB diagnosis. Secondary analysis was performed with these groups included. CI, confidence interval; TB, tuberculosis.

**Fig. 2. Log odds of documented tuberculosis diagnosis for negative (bottom line) and positive (top line) tuberculosis symptom screens across by age of screening.** Shaded regions represent 95% confidence intervals.
TB symptom screening for pediatric HIV

Vonasek et al.

77

likelihood ratio of 5.5 provides moderate increases in the post-test probability of disease [15]. TB symptom screening had an excellent NPV of 99.7%. However, the negative likelihood ratio of 0.44 only gives a 15–20% change in probability of no disease in the context of a negative screen [15], and over one-third of incident cases had a negative TB symptom screen, suggesting that a negative screen should be interpreted cautiously.

A strict case definition resulted in exclusion of 24% of those with documented TB disease from the primary analysis. Inclusion of these individuals in the secondary analysis resulted in small differences in accuracy estimates, most notably with a 6% decrease in estimated sensitivity. This decreased sensitivity may be explained by a higher proportion of negative screens being documented during visits with erroneously documented ‘TB disease’ (excluded from the primary analysis). Nonetheless, these differences between the primary and secondary analyses are unlikely to be clinically significant; it is unlikely that a 6% difference in sensitivity or a 0.5% difference in PPV would lead to different clinical decision-making. Across the age range screened in our study, no major differences in the predictive utility of TB symptom screening were noted. The log odds of TB being associated with the screening result was statistically different for negative and positive screens between ages of 7 and 17, and these differences were notably reduced in younger ages. This analysis suggests that symptom screening performed better in school-aged children and adolescents. However, even in these age groups, the difference in the log odds of disease between patients with a positive or negative symptom screen was so small that it is clinically insignificant.

A recent analysis examining the utility of ICF symptom screening in children less than 9 years of age living with HIV conducted in Soweto, South Africa demonstrated 57% sensitivity and 97% specificity [6], comparable with our findings. This study was strengthened by its prospective approach and use of a robust case definition for clinical diagnoses of TB. However, the external validity of this study has been questioned as it was conducted under strict research conditions at a single site that is part of a relatively well-resourced healthcare system compared with most TB/HIV high-burden settings [16]. In complement to these findings, our larger study includes adolescents, multiple countries in sub-Saharan Africa, and data collected under nonresearch conditions.

Improved performance of TB screening for CALHIV could enhance confidence in, and utilization of, screening results. When implemented by existing personnel, symptom screening is particularly attractive as it minimally strains resources while potentially affording cost-savings by streamlining the use of diagnostic assays like Xpert MTB/RIF, which healthcare systems in resource-limited, high-TB burden settings struggle to widely implement [17]. PPV of the current screening questions could be improved by utilizing a targeted approach to CALHIV with higher risk of TB disease, such as ART-naive CALHIV [6]. The symptom screening ICF strategy for adults has been shown to have lower sensitivity in those who are on antiretroviral therapy (ART) compared with those who are ART-naive [18]. In the era of ‘test and treat’, more people living with HIV are on ART compared with 2011 when these ICF strategies were developed. Unfortunately, this type of analysis was not possible with our dataset. Though a portion of the screening visits in this study involved patients that were ART-naive at the time, the majority of analyzed visits involved patients established on ART. Future research assessing variations in the TB symptom screen between CALHIV that are established in HIV care versus ART-naive, or based upon virologic control, could be valuable to develop more targeted approaches to screening.

Symptom screening could be done less frequently (e.g. quarterly for routine visits, or only at sick visits) to avoid ‘alarm fatigue’ from the abundance of false-positive screens, thereby improving the PPV and yield of the screen. Redefining a positive screen as the presence of two of four symptoms, rather than one of four, could further improve specificity, though likely at the expense of sensitivity. Marcy et al. [19] recently published a rigorous assessment of signs, symptoms, and diagnostic testing for the diagnosis of HIV-associated TB for a cohort of children under 14 years old from countries with a high burden of TB. Although this cohort only included children with presumptive TB based upon symptom screening for enrollment, and 60% of the cohort was ART-naive upon enrollment, the performance of the individual symptoms of the ICF symptom screen for predicting TB disease was assessed with a case–control sub-study. This revealed tremendous variability in the sensitivity and specificity of the four child ICF screening symptoms [19]. Available data for our study did not allow for assessment of individual or additive symptoms, or age-restricted subgroups, but further research in this area would also be informative. As more than one-third of those diagnosed with TB had negative screens, stand-alone use of the ICF four-symptom screen is not an acceptable rule-out strategy. There is very limited evidence to inform accurate screening strategies to support active TB case finding in CALHIV. Future research is needed to assess the utility of TB screening strategies in CALHIV that build off symptom screening by including novel strategies, such as blood C-reactive protein [5], automated chest radiograph interpretation algorithms [20], and microbiologic assays that rely on easily obtained specimens, such as urine or stool [21,22].

Although our findings are strengthened by the robust sample size and real-world context, there are limitations.
The retrospective approach is susceptible to detection bias as the clinicians were not blinded to screening result. TB symptom screens for a particular patient after diagnosis of an incident case were excluded, whereas patients without any diagnosis of TB disease have all screens included in the analysis. This difference led to some censoring of the data, which should be taken into consideration especially when interpreting the NPV and PPV. We only evaluated clinic visits with a documented TB symptom screen, and it is possible that TB symptoms screens were routinely performed but not accurately documented in the EMR, and thus not captured in our analysis. Evaluation of diagnostic and screening tools for TB in children, even under strict research conditions, is hampered by the lack of a sensitive reference standard [23–27]. Given that this data was collected from nonresearch settings, it is important to consider the reference standard here whenever interpreting the results. However, as described, these sites have relatively good access to chest radiography and microbiologic assays (e.g. Xpert MTB/RIF) and evaluation of presumptive cases is typically thorough.

Our evidence highlights the performance challenges of current TB symptom screening strategies among CAL-HIV, a population in whom the TB case detection gap is abysmally large. A more refined, targeted, and accurate TB screening approach coupled with better diagnostics is urgently needed to optimally identify TB disease in CAL/HIV.

Acknowledgements

The authors wish to thank all our patients and their caregivers for their involvement in this study, the staff at the BIPAI sites for their tireless care of these patients, and the monitoring and evaluation officers at the sites for their diligent management of this patient data over the years.

Contributors: B.V., A.K., J.B., D.D., H.H., and A.M. designed the study, T.D., D.D., and S.D. coordinated the data extraction from electronic medical records. D.D. conducted the statistical analysis. B.V. wrote the first manuscript draft. All authors contributed to the interpretation of the data, editing of the manuscript, and approved the final version of the manuscript.

Conflicts of interest

There are no conflicts of interest.

References

1. Jenkins HE, Yuen CM, Rodriguez CA, Nathavithanara RR, McLaughlin MM, Donald P, et al. Mortality in children diagnosed with tuberculosis: a systematic review and meta-analysis. Lancet Infect Dis 2017; 17:283–295.
2. Marais BJ, Ayres H, Graham SM, Godfrey-Faussett P. Screening and preventive therapy for tuberculosis. Clin Chest Med 2009; 30:827–846.
3. World Health Organization. Guidelines for intensified case-finding and isoniazid preventive therapy for people living with HIV in resource-constrained settings. Geneva, Switzerland: WHO; 2011.
4. Hermans S, Nasuua E, Van Less F, Byhoff E, Schwarz M, Hopelohman A, et al. Implementation and effect of intensified case finding on diagnosis of tuberculosis in a large urban HIV clinic in Uganda: a retrospective cohort study. BMC Public Health 2012; 12:674.
5. Yoon C, Dowdy DW, Esrail H, Macpherson P, Samuel G. Screening for tuberculosis: time to move beyond symptoms. Lancet Respir Med 2019; 7:202–204.
6. Savry S, Moultrie H, Van Rie A. Evaluation of the intensified tuberculosis case finding guidelines for children living with HIV. Int J Tuberc Lung Dis 2018; 22:1322–1328.
7. World Health Organization. Definitions and reporting framework for tuberculosis – 2013 revision. Geneva, Switzerland: WHO; 2014.
8. Wood S, Scheipl F. Estimate generalized additive mixed models via a version of function gamma from ‘mgcv’, using ‘lme4’ for estimation. R package version 0.2-6: 2020.
9. Kranzer K, Houben RM, Glynn JR, Bekker LG, Wood R, Lawn SD. Yield of HIV-associated tuberculosis during intensified case finding in resource-limited settings: a systematic review and meta-analysis. Lancet Infect Dis 2010; 10:q93–102.
10. World Health Organization. Global Tuberculosis Report 2018. Available at: http://www.who.int/tb/publications/global_report/en/ [Accessed 7 December 2018]
11. Chintu C, Mudenda V, Lucas S, Sumpi L, Singh M, Masuva D, et al. UNZA-UCLMS Project Paediatric Post-mortem Study Group. Lung diseases at necropsy in African children dying from respiratory illnesses: a descriptive necropsy study. Lancet 2002; 360:985–990.
12. Bates M, Shabetba M, Mudenda V, Chimoga C, Kembo J, Kabwe M, et al. Burden of respiratory tract infections at post mortem in Zambian children. BMC Med 2016; 14:1–9.
13. Ansari NA, Kombe AH, Kenyon TA, Hone NM, Tappero JW, Nyirenda ST, et al. Pathology and causes of death in a group of 128 predominantly HIV-positive patients in Botswana, 1997–1999. Int J Tuberc Lung Dis 2002; 6:53–63.
14. Dodd PI, Prendergast AJ, Bescott C, Kampmann B, Seddon JA. The impact of HIV and antiretroviral therapy on TB risk in children: A systematic review and meta-analysis. Thorax 2017; 72:559–575.
15. Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios. Lancet 2005; 365:1500–1505.
16. Kazembe PN. Intensive case finding. Int J Tuberc Lung Dis 2005; 12:1248–1247.
17. van’t Hoog AH, Cobelens F, Vassall A, van Kampen S, Dormann SE, Alland D, Eliner J. Optimal triage test characteristics to improve the cost-effectiveness of the Xpert MTB/RIF assay for TB diagnosis: a decision analysis. PLoS One 2013; 8:e82786.
18. Hamada Y, Lujan J, Schenkel K, Ford N, Getahun H. Sensitivity and specificity of WHO’s recommended four-symptom screening rule for tuberculosis in people living with HIV: a systematic review and meta-analysis. Lancet HIV 2018; 5:e515–e523.
19. Marcy O, Borand D, Uung V, Msellati P, Tejukem M, Truong Huu K, et al. ANRS 12229 PAANTHÉ 01 STUDY GROUP. A treatment-decision score for HIV-infected children with suspected tuberculosis. Pediatrics 2019; 144:e20182065.
20. Harris M, Qi A, Jeagal L, Torabi N, Menzies D, Korobitsyn A, et al. A systematic review of the diagnostic accuracy of artificial intelligence-based computer programs to analyze chest x-rays for pulmonary tuberculosis. PLoS One 2019; 14:e0221339.
21. World Health Organization. Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV. Policy update 2019. Geneva, Switzerland: WHO; 2019.
22. Atherton RR, Cresswell FV, Ellis J, Kitaka SB, Boulware DR. Xpert MTB/RIF Ultra for tuberculosis testing in children: a mini-review and commentary. Front Pediatr 2019; 7:34.
23. Nicol MP, Zar HJ. *Chapter 7: microbiological diagnosis of pulmonary tuberculosis in children*. In: Starke JR, Donald PR, editors. *Handbook of child & adolescent tuberculosis*. New York: Oxford University Press; 2016.

24. Detjen AK, DiNardo AR, Leyden J, Steingart KR, Menzies D, Schiller I, et al. *Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis in children: a systematic review and meta-analysis*. *Lancet Respir Med* 2015; 3:451–461.

25. DiNardo AR, Detjen A, Uster P, Ngo K, Bacha J, Mandalakas AM. *Culture is an imperfect and heterogeneous reference standard in pediatric tuberculosis*. *Tuberculosis (Edinb)* 2016; 101S:S105–S108.

26. Edwards DJ, Kitetele F, Rie A, XXX Van. *Agreement between clinical scoring systems used for the diagnosis of pediatric tuberculosis in the HIV era*. *Int J Tuberc Lung Dis* 2007; 11:263–269.

27. Du Toit G, Swingler G, Iloni K. *Observer variation in detecting lymphadenopathy on chest radiography*. *Int J Tuberc Lung Dis* 2002; 6:814–817.