scientific reports

OPEN

Diagnostic accuracy of screening and diagnostic tests used in a state-wide tuberculosis prevalence survey in India

Prathiksha Giridharan[1,2], Leeberk Raja Inbaraj[3,6✉], Asha Frederick[4], Sriram Selvaraju[1], Balaji Ramraj[3,6], Kannan Thiruvengadam[5], Bella Davaleenal Daniel[3] & Chandrasekaran Padmapriyadarsini[3]

Prevalence surveys and active case findings for tuberculosis (TB) typically use symptoms and Chest X-ray for screening and a diagnostic test for microbiological confirmation. We report the performance of screening and diagnostic tests from a state-wide prevalence survey in India. We screened 130,932 individuals using chest X-ray (CXR) and symptom checklist. Sputum samples were obtained from eligible participants and subjected to Xpert/MTB RIF, smear microscopy, and culture. Cough over two weeks had higher sensitivity (41.6%, 95% CI: 31.6–52.1) and lower specificity (72.8%, 95% CI: 72.1–73.5) among all the symptoms. The sensitivity and specificity of abnormal CXR were 86.4% (95% CI:77.9–92.5) and 42.12% (95% CI, 41.3–42.8), respectively. Xpert/MTB RIF in the reference laboratory had the highest sensitivity (96.55% 95% CI: 88.0-99.5), and smear microscopy had the highest specificity (99.7% 95% CI: 99.6–99.8). The combined sensitivity and specificity of Xpert MTB/RIF along symptom screening and CXR were 97.9% (95% CI:92.6–99.7) and 4.9% (95% CI:4.67–5.33), respectively. The mobile van Xpert MTB/RIF was found to be highly specific (99.3%). We recommend a diagnostic algorithm consisting of symptom screening and CXR followed by WHO-recommended rapid molecular assays (mWRD) for future prevalence surveys and active case finding.

Keywords Diagnostic test accuracy, Prevalence surveys, Active case finding, Xpert MTB/RIF

Globally, about 7.5 million people were newly diagnosed with tuberculosis (TB) and notified in the year 2022. Among them, 6.2 million (83%) had pulmonary TB. The proportion of people diagnosed with microbiologically confirmed TB improved between 2018 and 2021, from 55 to 63%; it remained at 63% in 2022[1]. One of the recommendations of the World Health Organization's (WHO) Global Task Force on TB Impact Measurement is to conduct national TB prevalence surveys in 22 priority countries as part of the End TB strategy. These survey results are essential for assessing progress towards global, regional and national targets for reductions in TB burden[2].

WHO also recommends conducting a systematic screening of the general population in regions with a TB prevalence of more than 0.5% and high-risk populations through active case finding (ACF)[3]. The cost implications of screening the healthy or high-risk population for TB infection and disease in low- and middle-income countries is a matter of concern. Simple, cheaper screening methods, such as symptom screening and chest X-rays followed by molecular diagnostic tests, are considered cost-effective strategies among the highest-risk groups[4]. ACF and prevalence surveys usually rely on symptom screening and chest X-ray for screening and diagnostic tests for microbiological confirmation. Sputum smear microscopy, WHO-recommended rapid molecular assays (mWRDs), and culture are the diagnostic tests that are most commonly used. Chada et al. argue that ACF exercise using molecular diagnostic tests would result in increased reporting of false positive cases and suggested evaluating the accuracy of molecular tests as a diagnostic tool in community settings[5]. Shewade et

[1]Department of Epidemiology, National Institute for Research in Tuberculosis Chennai, Chennai, India. [2]ICMR-National Instititute of Epidemiology, Chennai, India. [3]Department of Clinical Research, ICMR- National Institute for Research in Tuberculosis, Chennai, India. [4]Directorate of Medical and Rural Health Services, Chennai, Tamil Nadu, India. [5]Department of Statistics, ICMR- National Institute for Research in Tuberculosis, Chennai, India. [6]Faculty of Medical Research, Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, India. ✉email: leeberk2003@gmail.com

al. suggested that replacing smear microscopy with mWRDs may increase PPV to more than 90%, assuming the sensitivity of the mWRDs is similar in both ACF and PCF[6]. Most of the diagnostic accuracy measures of mWRDs are derived from passive case findings or studies done in healthcare facilities. While the evidence on the diagnostic accuracy of mWRDs in the ACF setting is limited, it could be extrapolated from national prevalence surveys. The best estimates of the sensitivity and specificity of Xpert MTB/RIF in the general population in the national prevalence surveys were 69% and 99%, respectively[7].

Furthermore, the TB program would benefit from investigating the performance of the screening tests that are followed by a diagnostic test in the sequential screening strategy. Myanmar national prevalence survey estimated screening for any symptoms and signs suggestive of TB had a sensitivity of 59.8% (95% CI:54.1–65.3) and a specificity of 67.2% (95% CI: 66.7–67.6)[8]. Kenya National Tuberculosis survey revealed that the sensitivity of CXR being interpreted by a medical officer as "suggestive of TB" for bacteriologically-confirmed TB was 43.7% (95% Credible interval (CrI): 23.8–66.4%) and specificity was 89.2% (95% CrI: 89.0–89.6%)[9]. In the context of limited evidence on the performance of screening and diagnostic tests for TB disease detection in the community setting, we report the diagnostic accuracy of the screening and diagnostic tests used in the Tamil Nadu TB prevalence survey (TNTBPS). Our findings from this survey might help the TB program managers, especially in the high TB burden setting, to decide on the appropriate diagnostic algorithm to be used for ACF and prevalence surveys.

Results

We screened 130,932 individuals who consented to participate in the survey, and all participants were screened at least by one method. Among the screened population, 130,914 (99.9%) underwent symptom screening, and 125,870 (96.1%) completed both symptom screening and CXR examination. Of 130,932, 20,086 (15.34%) were found to be eligible for sputum collection, and eventually, 18,669 (92.9%) provided sputum for testing. The reasons for the inability to obtain sputum from all the eligible participants were the inability to produce sputum, unwillingness, and survey loss to follow-up. The first sputum sample was obtained from 18,654 (92.8%), and the second sputum sample was collected from 18,255 (90.8%) participants. Among the 18,669 samples, 17,184 participants had valid results of MTB/RIF, smear microscopy, and MGIT. We have removed 106 participants who were on treatment for tuberculosis and considered 17,078 for the final analysis of diagnostic accuracy. Of the 17,078 samples processed in the survey van for Xpert MTB/RIF, 185 (1.08%) tested positive for MTB (Fig. 1). In the reference laboratory's smear microscopy, 89 (0.52%) samples were positive and 96 (0.56%) were positive in MGIT. The participants included predominantly females (50.2%), and almost one-tenth of them gave a history of previous treatment (9.1%) for TB (Table 1).

Diagnostic accuracy of screening tests

The sensitivities of the symptom screening checklist ranged from 3.1 to 41.6%, while the specificities ranged from 72.8 to 98.6%. Among the symptoms, cough of more than two weeks had higher sensitivity (41.6%, 95% CI:31.6–52.1) and lower specificity (72.8%, 95% CI:72.1–73.5) among all symptoms (Table 2). Fever for 14 days or more had higher specificity (98.60% 95% CI: 98.4–98.7). The presence of any one of the symptoms suggestive of TB had a sensitivity of 55.2% (95% CI: 44.7–65.3) and a specificity of 50.9% (95% CI: 50.1–51.6). Among all the screening tests, the sensitivity of an abnormal chest X-ray was the highest (86.46%, 95% CI: 77.9–92.5), while its specificity was lower (42.12%, 95% CI: 41.3–42.8) compared to symptom screening. When CXR was combined with symptom screening, the sensitivity rose by nearly 11% points (97.9%, 95% CI: 92.6–99.7), but the specificity significantly decreased (4.99%, 95% CI: 4.67–5.33).

Diagnostic accuracy of diagnostic tests

Among the diagnostic tests, Xpert/MTB RIF in the reference laboratory had the highest sensitivity (96.5%, 95% CI:88–99.5) (Table 3). Smear microscopy had the highest specificity (99.7%, 95% CI: 99.6–99.8), followed by Xpert assay in the mobile van (99.3%, 95% CI: 99.1–99.4). The sensitivity of Xpert MTB/RIF in the mobile van was 71.8% (95% CI: 61.7–80.5), and the sensitivity of smear microscopy was the lowest (53.13%, 95% CI: 42.6–63.3) among the diagnostic tests used in the survey. The kappa coefficient for smear microscopy and MGIT between sputum 1 and 2 were 0.64 (SD- 0.077, substantial agreement) and 0.592 (SD-0.077, moderate agreement), respectively.

Diagnostic accuracy of a combination of screening and diagnostic tests

We analyzed the combination of screening and diagnostic tests used in the survey. Xpert MTB/RIF with symptom screening alone yielded a sensitivity of 84.38% (95% CI: 75.5–90.9) and specificity of 50.6% (95% CI:49.9–51.4) (Table 4). When Xpert MTB/RIF was combined with CXR and demonstrated a sensitivity of 93.75% (95% CI: 86.89–97.67) and specificity of 42.50% (95% CI: 41.76–43.25) in the survey van. The sensitivity (90.63%) was higher in the reference laboratory with almost similar specificity (42.60%. When CXR was added to symptom screening along with Xpert/MTB RIF, the sensitivity increased almost by 15% points (98.9% 95% CI: 94.3–99.9) with the decrease in specificity (4.98%, 95% CI: 4.6–5.3). This phenomenon was observed even with adding smear microscopy, with symptoms screening and CXR. However, the incremental yield in sensitivity was slightly lower for the Xpert assay (97.9%, 95% CI: 92.6–99.7) in the reference laboratory compared to the mobile van.

Discussion

Cough has been one of the most sensitive symptoms that is suggestive of tuberculosis in passive as well as ACF for TB. WHO's screening guidelines reported a sensitivity of 42% for cough among HIV-negative individuals, and this finding corroborates with our survey (41.6%)[3]. However, this was lower than the prevalence survey
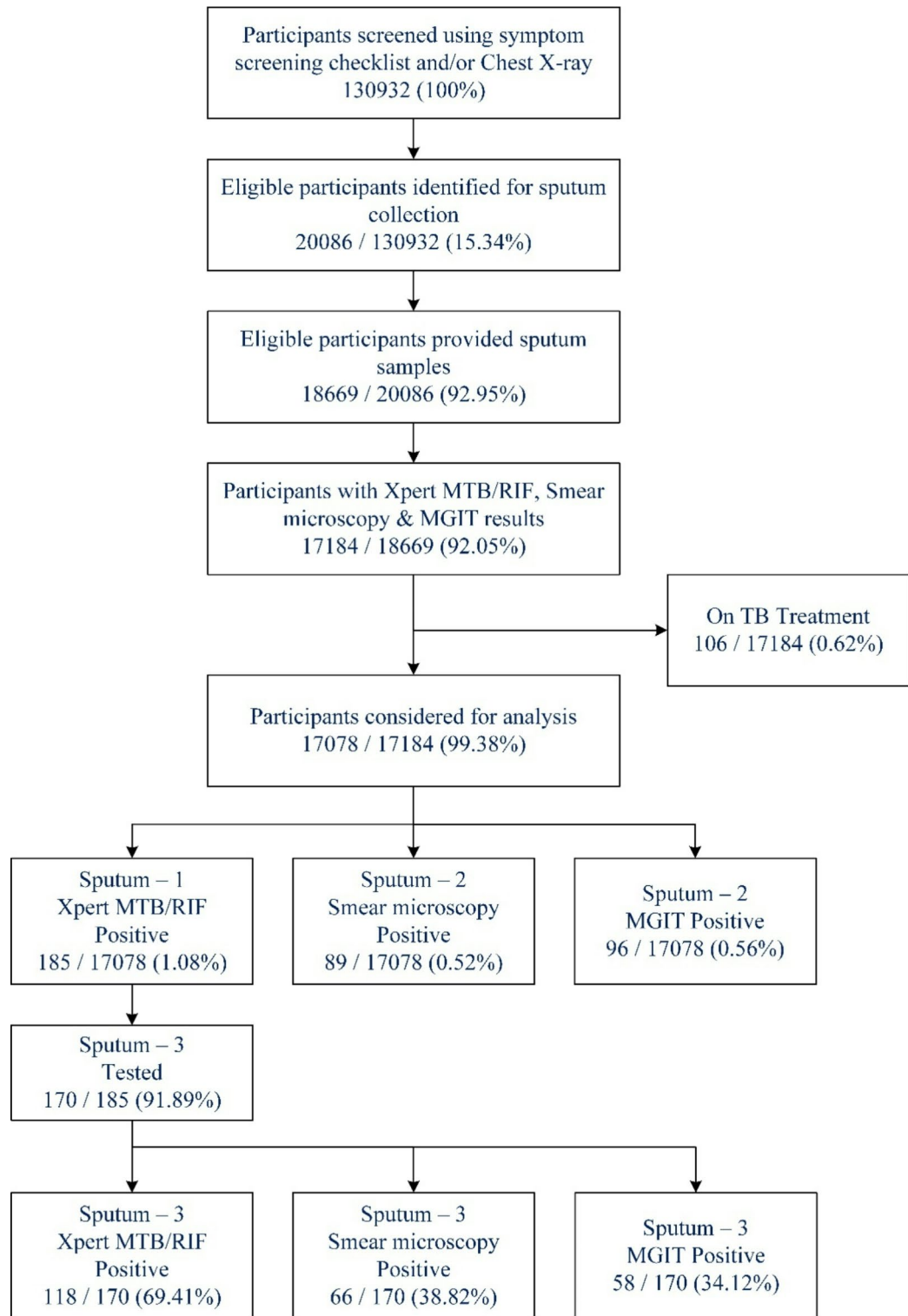
Fig. 1. Flow diagram showing number of participants screened, eligible for sputum, sputum samples tested and their results.

from Kenya [52 (95% CI: 41–63)] which was done between 2005 and 07[10]. The specificity (72.8%) of cough in our survey was also lower than that of the Kenyan survey [89% (95% CI: 88–90)[10]. The Kenyan TB prevalence survey reported a sensitivity of 90% (95% CI:84–95) and a specificity of 32% (95% CI: 30–34) for the presence of any one symptom suggestive of TB. While we estimated a lower sensitivity (55.2%) than the Kenyan survey, our specificity was higher (50.9%) for any one symptom suggestive of TB. This is likely attributable to the fact that the survey conducted in Kenya included a substantial number of individuals who were HIV positive or

Variables	N (%)
Age (in years)	
15–24	1131 (6.6%)
25–34	1446 (8.5%)
35–44	2564 (15.0%)
45–54	3438 (20.1%)
55–64	4061 (23.8%)
65+	4438 (26.0%)
Gender	
Female	8568 (50.2%)
Male	8510 (49.8%)
Past History of TB	
No	15,517 (90.9%)
Yes	1561 (9.1%)
Self-Reported Smoking Status	
Non-Smoker	14,069 (82.4%)
Smoker	3009 (17.6%)
Self-Reported Alcoholic Status	
Non-Alcoholic	13,424 (78.6%)
Alcoholic	3654 (21.4%)

Table 1. Demographic characteristics of study participants. Categorical variables were presented as frequency (percentages)

had an unknown HIV status. Our sensitivity estimation for the presence of any one symptom is very similar to the survey conducted in Myanmar (59.8%), while the specificity (50.9%) is lower than that of the Myanmar survey (67.2%)[8]. A Cochrane review of 31 studies done among participants who were screened for tuberculosis estimated sensitivity of 70.6% (95% CI:61.7–78.2%) and specificity of 65.1% (95% CI:53.3–75.4%) for any tuberculosis symptoms. Symptom screening is the most simple screening tool that can be used even in limited resource settings. Though this tool is considered to have low accuracy, we found the specificity of symptom screening was higher than that of CXR[11].

The sensitivity (87.25%) of CXR in our survey was almost similar to the Kenyan survey. However, the specificity (90.62%) was lower. CXR has been widely used in ACF and prevalence surveys and is known to be a screening tool with high sensitivity[10]. The aforesaid Cochrane review also included 19 studies and estimated sensitivity and specificity of 84.8% (95% CI:76.7–90.4) and 95.6% (95% CI: 92.6–97.4), respectively, for CXR abnormalities suggestive of TB[11]. Though CXR sensitivity in our survey was higher, the specificity was lower than in this review. When symptom screening was combined with CXR, it increased the sensitivity (98.04%) significantly at the cost of reduction in specificity. This indicates that CXR and symptom screening combined can significantly reduce false negatives in the screening population. Though CXR is a good screening test, it has a few limitations, such as the necessity for radiation-shielded vehicles, the higher cost associated with acquiring and upkeeping the vehicles and X-ray equipment, the availability of technicians; the crucial requirement for medical officers or radiologists to interpret CXR results; the inconsistencies in reporting between different observers and even within the same observer; and the exposure of apparently healthy individuals to radiation. However, these limitations and barriers may be surmounted by the recent advancements in portable X-ray machines and artificial intelligence to facilitate reporting. Given that our results unequivocally indicate that it is beneficial to allocate resources toward advanced CXR technology in both ACF and prevalence surveys.

We estimated a sensitivity of 71.88% (95% CI: 61.7–80.5) for Xpert MTB/RIF performed in the survey van. This is higher than the sensitivity of Xpert MTB/RIF in the recent prevalence surveys conducted in Kenya (69%), the Philippines (69%) and Vietnam (68%)[7]. However, our sensitivity estimate is lower than that of Bangladesh's 84% (95% CI: 78–84) prevalence survey. The specificities of Xpert MTB/RIF from these surveys corroborate our findings. The sensitivity of the molecular test in the reference laboratory was significantly higher than the survey van [96.55% (95% CI: 88.0 −99.5)]. We could postulate several reasons for the decrease in sensitivity, such as temperature in the field setting, and environmental factors. The pooled sensitivity of Xpert Ultra (78%, 95% CI: 69–84%) was higher than Xpert MTB/RIF (73%, 95% CI: 62–82%) from the recent prevalence surveys (South Africa, Myanmar, Lesotho, and Zambia) conducted between 2017 and 2019[7]. Future surveys should also consider WHO-recommended low complexity automated NAATs such as Truenat MTB Plus and Xpert Ultra to increase the yield in the survey as they have shown better performance in the healthcare setting and passive case finding. It is worthwhile to note that though smear microscopy had the lowest sensitivity (53.13% 95% CI: 42.6–63.9), it had the highest specificity (99.78% 95% CI: 99.6–99.8) among all the diagnostic tests. This implies that smear microscopy could still be used to confirm a diagnosis when molecular tests are unavailable in resource-limited settings.

When we combined our screening test and diagnostic tests, there was a significant increase in sensitivity with a reduction in specificity. However, our screening and diagnostic approach yielded a negative predictive value

	MGIT + ve	MGIT -ve	Measures (95% Confidence interval)	
Cough (*n* = 17078)				
			Sensitivity	41.67% (31.68, 52.18)
≥ 2 weeks	40	4,614	Specificity	72.83% (72.15, 73.50)
	41.7%	27.2%	PPV	0.86% (0.61, 1.17)
< 2 Weeks	56	12,368	NPV	99.55% (99.42, 99.66)
	58.3%	72.8%	Accuracy	72.65% (71.98, 73.32)
			PLR: 1.53, NLR: 0.80	
Fever (n = 17078)				
			Sensitivity	3.13% (0.65, 8.86)
≥ 2 weeks	3	237	Specificity	98.60% (98.42, 98.78)
	3.1%	1.4%	PPV	1.25% (0.26, 3.61)
< 2 Weeks	93	16,745	NPV	99.45% (99.32, 99.55)
	96.9%	98.6%	Accuracy	98.07% (97.85, 98.27)
			PLR: 2.24, NLR: 0.98	
Weight Loss (n = 17078)				
			Sensitivity	10.42% (5.11, 18.32)
Yes	10	1,835	Specificity	89.19% (88.72, 89.66)
	10.4%	10.8%	PPV	0.54% (0.26, 0.99)
No	86	15,147	NPV	99.44% (99.30, 99.55)
	89.6%	89.2%	Accuracy	88.75% (88.27, 89.22)
			PLR: 0.96, NLR: 1.00	
Blood in Sputum in last 6 months (n = 17078)				
			Sensitivity	6.25% (2.33, 13.11)
Yes	6	1,027	Specificity	93.95% (93.58, 94.31)
	6.3%	6.0%	PPV	0.58% (0.21, 1.26)
No	90	15,955	NPV	99.44% (99.31, 99.55)
	93.8%	94.0%	Accuracy	93.46% (93.08, 93.83)
			PLR: 1.03, NLR: 1.00	
Chest Pain in last one month (n = 17078)				
			Sensitivity	18.75% (11.51, 28.00)
Yes	18	3,636	Specificity	78.59% (77.96, 79.20)
	18.8%	21.4%	PPV	0.49% (0.29, 0.78)
No	78	13,346	NPV	99.42% (99.28, 99.54)
	81.3%	78.6%	Accuracy	78.25% (77.63, 78.87)
			PLR: 0.88, NLR: 1.03	
Past history of TB Treatment (n = 17078)				
			Sensitivity	12.50% (6.63, 20.82)
Yes	12	1,549	Specificity	90.88% (90.44, 91.31)
	12.5%	9.1%	PPV	0.77% (0.40, 1.34)
No	84	15,433	NPV	99.46% (99.33, 99.57)
	87.5%	90.9%	Accuracy	90.44% (89.99, 90.88)
			PLR: 1.37, NLR: 0.96	
Any one of the above screening positive (n = 17078)				
			Sensitivity	55.21% (44.71, 65.37)
Yes	53	8,332	Specificity	50.94% (50.18, 51.69)
	55.2%	49.1%	PPV	0.63% (0.47, 0.83)
No	43	8,650	NPV	99.51% (99.33, 99.64)
	44.8%	50.9%	Accuracy	50.96% (50.21, 51.71)
			PLR: 1.13, NLR: 0.88	
Chest X-ray Findings (n = 16925)				
			Sensitivity	86.46% (77.96, 92.59)
Abnormal	83	9,741	Specificity	42.12% (41.37, 42.87)
	86.5%	57.9%	PPV	0.84% (0.67, 1.05)
Continued				
Normal	13	7,088	NPV	99.82% (99.69, 99.90)
	13.5%	42.1%	Accuracy	42.37% (41.62, 43.12)

5

	MGIT + ve	MGIT -ve		Measures (95% Confidence interval)
Cough (n = 17078)				
				PLR: 1.49, NLR: 0.32
Eligible for Sputum by Screening + Abnormal chest findings in CXR (n = 17078)				
			Sensitivity	97.92% (92.68, 99.75)
Yes	94	16,135	Specificity	4.99% (4.67, 5.33)
	97.9%	95.0%	PPV	0.58% (0.47, 0.71)
No	2	847	NPV	99.76% (99.15, 99.97)
	2.1%	5.0%	Accuracy	5.51% (5.17, 5.86)
				PLR: 1.03, NLR: 0.42

Table 2. Diagnostic accuracy of screening tests. Participants who were on TB treatment were excluded

	MGIT + ve	MGIT -ve		Measures (95% Confidence interval)
Xpert MTB/RIF (Van; sample-1) vs. MGIT (n = 17078)				
			Sensitivity	71.88% (61.78, 80.58)
Xpert + ve	69	116	Specificity	99.32% (99.18, 99.44)
	71.9%	0.7%	PPV	37.30% (30.31, 44.70)
Xpert-ve	27	16,866	NPV	99.84% (99.77, 99.89)
	28.1%	99.3%	Accuracy	99.16% (99.01, 99.29)
				PLR: 105.22, NLR: 0.28
AFB Smear (Van; sample − 2) vs. MGIT (n = 17078)				
			Sensitivity	53.13% (42.66, 63.39)
Smear + ve	51	38	Specificity	99.78% (99.69, 99.84)
	53.1%	0.2%	PPV	57.30% (46.37, 67.74)
Smear-ve	45	16,944	NPV	99.74% (99.65, 99.81)
	46.9%	99.8%	Accuracy	99.51% (99.40, 99.61)
				PLR: 237.41, NLR: 0.47
Xpert MTB/RIF (IRL; sample − 3) vs. MGIT (n = 170)				
			Sensitivity	96.55% (88.09, 99.58)
Yes	56	62	Specificity	44.64% (35.24, 54.33)
	96.6%	55.4%	PPV	47.46% (38.19, 56.85)
No	2	50	NPV	96.15% (86.79, 99.53)
	3.4%	44.6%	Accuracy	62.35% (54.61, 69.66)
				PLR: 1.74, NLR: 0.08
AFB Smear (IRL; Sample-3) vs. MGIT (n = 170)				
			Sensitivity	67.24% (53.66, 78.99)
Yes	39	27	Specificity	75.89% (66.90, 83.47)
	67.2%	24.1%	PPV	59.09% (46.29, 71.05)
No	19	85	NPV	81.73% (72.95, 88.63)
	32.8%	75.9%	Accuracy	72.94% (65.61, 79.46)
				PLR: 2.79, NLR: 0.43

Table 3. Diagnostic accuracy of diagnostic tests (Xpert MTB/RIF, smear miscroscopy). Participants who were on TB treatment were excluded

of 99.8%. Our participants underwent highly sensitive screening tests followed by a highly specific confirmatory test. CXR yielded a significant proportion of false positive results, which were later eliminated by the Xpert MTB/RIF. The possible reasons could be that the survey was done during and immediately after COVID-19, and the CXR was interpreted by both trained medical officers and specialists. The primary objective of the prevalence survey is to identify the true prevalence in the community, including cases that were not routinely detected by the passive case finding through healthcare system. This will enable us to estimate the prevalence notification gap, which is an initial step of the TB care cascade and an indicator of the efficiency of the TB management system in the state. Conducting an initial screening combining symptoms and chest X-rays may significantly reduce the likelihood of false negative results and enable participants to proceed with the specified diagnostic

		MGIT +ve		MGIT -ve	Measures (95% Confidence interval)	
Symptom screening + Xpert MTB/RIF (Van ; sample 2) (n = 17078)						
					Sensitivity	84.38% (75.54, 90.98)
Yes		81		8,374	Specificity	50.69% (49.93, 51.44)
		84.4%		49.3%	PPV	0.96% (0.76, 1.19)
No		15		8,608	NPV	99.83% (99.71, 99.90)
		15.6%		50.7%	Accuracy	50.88% (50.13, 51.63)
					PLR: 1.71, NLR: 0.31	
CXR + Xpert MTB/RIF (Van; sample 2) (n = 17078)						
					Sensitivity	93.75% (86.89, 97.67)
Yes		90		9,764	Specificity	42.50% (41.76, 43.25)
		93.8%		57.5%	PPV	0.91% (0.74, 1.12)
No		6		7,218	NPV	99.92% (99.82, 99.97)
		6.3%		42.5%	Accuracy	42.79% (42.05, 43.54)
					PLR: 1.63, NLR: 0.15	
Symptom screening + CXR + Xpert MTB/RIF (Van; sample 2) (n = 17078)						
					Sensitivity	98.96% (94.33, 99.97)
Yes		95		16,136	Specificity	4.98% (4.66, 5.32)
		99.0%		95.0%	PPV	0.59% (0.47, 0.72)
No		1		846	NPV	99.88% (99.34, 100.00)
		1.0%		5.0%	Accuracy	5.51% (5.17, 5.86)
					PLR: 1.04, NLR: 0.21	
Symptom screening + Xpert MTB/RIF (IRL; sample 3) (n = 17078)						
					Sensitivity	81.25% (72.00, 88.49)
Yes		78		8,354	Specificity	50.81% (50.05, 51.56)
		81.3%		49.2%	PPV	0.93% (0.73, 1.15)
No		18		8,628	NPV	99.79% (99.67, 99.88)
		18.8%		50.8%	Accuracy	50.98% (50.23, 51.73)
					PLR: 1.65, NLR: 0.37	
CXR + Xpert MTB/RIF (IRL; sample 3) (n = 17078)						
					Sensitivity	90.63% (82.95, 95.62)
Yes		87		9,748	Specificity	42.60% (41.85, 43.35)
		90.6%		57.4%	PPV	0.88% (0.71, 1.09)
No		9		7,234	NPV	99.88% (99.76, 99.94)
		9.4%		42.6%	Accuracy	42.87% (42.12, 43.61)
					PLR: 1.58, NLR: 0.22	
Symptom screening + CXR + Xpert MTB/RIF (IRL ; sample 3) (n = 17078)						
					Sensitivity	97.92% (92.68, 99.75)
Yes		94		16,135	Specificity	4.99% (4.67, 5.33)
		97.9%		95.0%	PPV	0.58% (0.47, 0.71)
No		2		847	NPV	99.76% (99.15, 99.97)
		2.1%		5.0%	Accuracy	5.51% (5.17, 5.86)
					PLR: 1.03, NLR: 0.42	
Symptom screening + smear microscopy (Van; sample 2) (n = 17078)						
					Sensitivity	75.00% (65.12, 83.28)
Yes		72		8,348	Specificity	50.84% (50.09, 51.60)
		75.0%		49.2%	PPV	0.86% (0.67, 1.08)
No		24		8,634	NPV	99.72% (99.59, 99.82)
		25.0%		50.8%	Accuracy	50.98% (50.23, 51.73)
					PLR: 1.53, NLR: 0.49	
CXR + Smear microscopy (Van; sample 2) (n = 17078)						
					Sensitivity	91.67% (84.24, 96.33)
Yes		88		9,751	Specificity	42.58% (41.84, 43.33)
		91.7%		57.4%	PPV	0.89% (0.72, 1.10)
No		8		7,231	NPV	99.89% (99.78, 99.95)
		8.3%		42.6%	Accuracy	42.86% (42.11, 43.60)
					PLR: 1.60, NLR: 0.20	
Continued						

	MGIT +ve	MGIT -ve	Measures (95% Confidence interval)	
Symptom screening + Xpert MTB/RIF (Van ; sample 2) (*n* = 17078)				
Symptom screening + CXR + smear microscopy (Van; sample 2) (n = 17078)				
			Sensitivity	98.96% (94.33, 99.97)
Yes	95	16,137	Specificity	4.98% (4.65, 5.31)
	99.0%	95.0%	PPV	0.59% (0.47, 0.71)
No	1	845	NPV	99.88% (99.34, 100.00)
	1.0%	5.0%	Accuracy	5.50% (5.17, 5.86)
				PLR: 1.04, NLR: 0.21
Past history of TB + Xpert MTB/RIF (van; sample 2) (n = 17078)				
			Sensitivity	72.92% (62.89, 81.48)
Yes	70	1,642	Specificity	90.33% (89.88, 90.77)
	72.9%	9.7%	PPV	4.09% (3.20, 5.14)
No	26	15,340	NPV	99.83% (99.75, 99.89)
	27.1%	90.3%	Accuracy	90.23% (89.78, 90.67)
				PLR: 7.54, NLR: 0.30
Past history of TB + Xpert MTB/RIF (van; sample 3) (n = 17078)				
			Sensitivity	62.50% (52.03, 72.18)
Abnormal	60	1,598	Specificity	90.59% (90.14, 91.03)
	62.5%	9.4%	PPV	3.62% (2.77, 4.63)
Normal	36	15,384	NPV	99.77% (99.68, 99.84)
	37.5%	90.6%	Accuracy	90.43% (89.98, 90.87)
				PLR: 6.64, NLR: 0.41

Table 4. Diagnostic accuracy of screening and diagnostic tests. Participants who were on TB Treatment were excluded; Symptoms includes either of the cough, fever, weight loss, blood in sputum, chest pain, and past history of TB

tests. We found that the True positive rate was significantly higher when all the tests were combined. We also noted similar estimates when the molecular test was replaced with smear microscopy.

Methodology
Tamil Nadu is a southern state in India with 33 administrative districts and more than a population of 60 million. The survey covered 180 clusters in all the districts with a multistage cluster sampling design with a target sample size of 144,000 and screened 130,932 individuals from the 143,005 eligible population[12]. The number of clusters in each district was allocated according to the population i.e. population proportionate to size.

Participants and screening procedures
We included 15 years and older participants in the selected village or urban census enumeration block. We excluded hospitalised residents, institutional populations, and sick or bedridden individuals. All the eligible participants initially underwent screening by a symptom checklist followed by a digital chest X-ray (PA view) in the mobile survey van. Trained survey staff screened the population with a symptom checklist that consisted of fever, cough, weight loss, haemoptysis, and chest pain in the last month prior to the survey. Participants with the presence of any of the symptoms, abnormal chest X-ray, or a history of current or past anti-TB treatment were defined to be eligible for sputum testing and asked to give sputum for examination.

Study procedures
The mobile survey van consists of an X-ray unit, Xpert/MTB RIF machines, a refrigerator, and a laptop. Eligible participants provided 3 to 5 ml of sputum for testing. The first sample was processed and analyzed in the mobile van used for the survey using Xpert MTB/RIF to detect MTB and rifampicin resistance. Subsequently, a second sputum sample was collected within 24 h and transported within 48 h to a reference laboratory which was identified before the survey. The second sample was subjected to AFB smear microscopy, liquid culture (Mycobacteria Growth Indicator Tube) and drug susceptibility testing (DST) in the reference laboratory located in Chennai and Madurai in Tamil Nadu, India. If MTB was detected by Xpert MTB/RIF in the survey van from the first sample, a third sample was collected and transported to the reference laboratory for Xpert MTB/RIF, smear microscopy, liquid culture and drug susceptibility testing (DST). Acid-fact bacilli smears were examined with fluorescence microscopy. Xpert MTB/RIF (Cepheid. Inc) was performed in accordance with the manufacturer's instructions. Xpert MTB/RIF results are automatically generated (i.e., there is a single threshold), and the user is provided with printable test results. Indeterminate results were reported after repeating the test. Xpert MTB/RIF results done in the survey van were not shared with the reference laboratory. The samples were de-contaminated at the reference laboratory and inoculated into MGIT960 (Becton Dickinson, Sparks, MD, USA), a liquid culture method for the detection of mycobacteria. Subsequently, speciation was done for all

the positive cultures and was confirmed using smear microscopy and TBc Identification test (TBcID, Becton Dickinson, Sparks, USA) as *M. tuberculosis*.

Quality assurance

The trained technicians performed all the laboratory tests in the study according to the standard operating procedures (SOPs). Temperature monitors were used to maintain the refrigerator for reagents storage at $2^{o}-8^{o}C$ and the Xpert cabin at $20^{o} - 25^{o}C$ in the mobile vans. Using a specialized transporting system, it was ensured that the samples reached the reference laboratories within 24–28 h. Our quality assurance team consists of microbiologists, scientists from ICMR-NIRT, and district TB officers visited each participating laboratory and ensured the procedures were done as per the SOPs. National TB elimination program (NTEP) recommendations were followed for the quality assurance of smear, Xpert/MTB, and culture in reference laboratories. All the medical officers, X-ray technicians, and teleradiologists were trained in all the SOPs. If the medical officer in the field carried out the spot quality check of the X-ray image and if the image was not acceptable, the Chest X-ray was repeated again. Two teleradiologists reported all the CXRs, and if there was a discrepancy between the radiologists, a third radiologist provided the final report.

Data entry

Field data was collected electronically using a customized Android application, while data from reference laboratories and the teleradiology panel was entered into a customized online web application designed for the survey. The survey data was monitored periodically, synced and stored in the server at ICMR-NIRT, Chennai. Data cleaning was done to remove duplication, outliers, and misclassification of variables before the analysis.

Data analysis

All the statistical analysis was done using Stata 16.0 (Stata Corporation, College Station, TX, USA). Demographic characteristics were analyzed and presented as numbers and percentages. We excluded those who were currently on TB treatment for the analysis of diagnostic accuracy. Sensitivity, specificity, positive predictive values (PPVs), negative predictive values (NPVs), positive likelihood ratio (PLR) and negative likelihood ratio (NLR) were calculated and presented with 95% confidence intervals. We calculated the diagnostic accuracy of screening (symptom screening and CXR) and diagnostic (Xpert MTB/RIF and smear microscopy) against MGIT as a reference standard. We evaluated the accuracy of a combination of diagnostic tests and symptom screening. A positive result for a diagnostic test or the presence of any one symptom was considered positive and compared to the MGIT. A positive MGIT result was considered as disease positive for our analysis.

Strength and limtations

Our survey had several strengths. We strictly adhered to the design and methodology recommended by the WHO. The staff were trained in the screening and diagnostic procedures and they were closely monitored for adherence to the SOPs of the survey. We had difficulty in obtaining good quality sputum from the participants. The survey included the apparently healthy individuals who could come to the survey site. We could have missed individuals who were sick and bed-ridden at home or unable to stand for CXR.

Conclusion

Symptom screening and CXR were the highly sensitive screening tool that can be used in the prevalence surveys. mWRDs (Xpert/MTB RIF) was found to be highly specific even in the prevalence surveys. We recommend a diagnostic algorithm consisting of symptom screening, CXR followed by mWRDs for the future prevalence surveys and active case finding.

Data availability

A subset of the key anonymised individual participant data collected during the study, along with a data dictionary, is available upon request to the corresponding author after approval of a proposal with a signed data access agreement.

References

1. WHO Global Tuberculosis. report [internet] 2023 [cited on 2024 August 25] Available from: https://iris.who.int/bitstream/handle/10665/373828/9789240083851-eng.pdf?sequence=1
2. Global Task Force on TB Impact Measurement [internet] [cited on. August 25] (2024). Available from: https://www.who.int/groups/global-task-force-on-tb-impact-measurement
3. WHO consolidated guidelines on tuberculosis. Module.2 Systematic screening for tuberculosis disease [internet] [cited on 2024 August 25] Available from: https://iris.who.int/bitstream/handle/10665/340255/9789240022676-eng.pdf
4. Alsdurf, H. et al. Tuberculosis screening costs and cost-effectiveness in high-risk groups: a systematic review. *BMC Infect. Dis.* **21**, 935. https://doi.org/10.1186/s12879-021-06633-3 (2021).
5. Chadha, V. K. & Praseeja, P. Active tuberculosis case finding in India - The way forward. *Indian J. Tuberc.* **66** (1), 170–177. https://doi.org/10.1016/j.ijtb.2018.05.014 (2019). Epub 2018 Jun 1. PMID: 30878064.
6. Shewade, H. D., Satyanarayana, S. & Kumar, A. M. Does active case finding for tuberculosis generate more false-positives compared to passive case finding in India? *Indian J. Tuberc.* **68** (3), 396–399 (2021). Epub 2020 Sep 17. PMID: 34099207.
7. WHO.National tuberculosis. prevalence surveys: what diagnostic algorithms should be used in future? [internet] [cited on 2024 August 25]. Available from : https://iris.who.int/bitstream/handle/10665/367909/9789240073913-eng.pdf?sequence=1

8. Ko Htet, K. K., Chongsuvivatwong, V. & Aung, S. T. Sensitivity and specificity of tuberculosis signs and symptoms screening and adjunct role of social pathology characteristics in predicting bacteriologically confirmed tuberculosis in Myanmar. *Trop. Med. Health*. 49. https://doi.org/10.1186/s41182-020-00292-x (2021).

9. Mungai, B., Ong'angò, J., Ku, C. C. & Henrion, M. Y. R. Accuracy of computer-aided chest X-ray in community-based tuberculosis screening: lessons from the 2016 Kenya National tuberculosis prevalence survey. *PLOS Glob Public. Health*. **2** (11), e0001272. https://doi.org/10.1371/journal.pgph.0001272 (2022).

10. van't Hoog, A. H. et al. Screening strategies for tuberculosis prevalence surveys: the valueof chest radiography and symptoms. *PLoS One* **7**(7), e38691. https://doi.org/10.1371/journal.pone.0038691 (2012).

11. Prathiksha, G. et al. Programmatic implications of a sub-national TB prevalence survey in India. *Int. J. Tuberc Lung Dis.* **28** (7), 348–353 (2024).

12. van't Hoog, A. et al. Symptom- and chest-radiography screening for active pulmonary tuberculosis in HIV-negative adults and adults with unknown HIV status. *Cochrane Database Syst. Rev. [Internet]*. **3**(3), 1–261 (2022). Available from: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD010890.pub2/full.

Acknowledgements

Author contributions

PG was the prinicipal investigator of the prevalence survey. LR and PG conceived this manuscirpt, inetrpreted the analysis and wrote the first draft. KT performed the statistcial analysis. AF supervised the survey and critically reveiwed the manuscript. RB and BD wrote gave inputs and critically reviewed the manuscript. SS and CP critically reviwed the manuscript. All authors reviewed and accepted the final version of the manuscript.

Funding

Declarations

Ethics approval and consent to participate

The study was approved by the ethics committee of the ICMR- National Institute for Research in Tuberculosis (017/NIRT-IEC/2021) and informed consent was obtained from all the participants, and pregnant women were excluded from being subjected to chest X-rays. The prevalence survey was performed in accordance with the WHO's guidelines for tuberculosis prevalence surveys.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.R.I.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.