The Union
The Official Journal of The International Union Against Tuberculosis and Lung Disease
The International Journal of Tuberculosis and Lung Disease (IJTLD)
ISSN 1027-3719

PROOF READING INSTRUCTIONS FOR AUTHORS

Attached are page proofs of your article, which is to appear in *The International Journal of Tuberculosis and Lung Disease* (IJTLD).

Please note, any questions that arose during copy editing is flagged in the margins and will appear in full at the end of the article.

Please address any typesetting errors promptly and return your comments within 72 hours by e-mail to: journal@theunion.org – a failure to reply within 72 hours may delay the publication of your article.

If you have any questions, please contact the Editorial Office at journal@theunion.org

Int J Tuberc Lung Dis 30(4):1–9

REVIEW

The diagnostic accuracy of interferon-gamma release assay for TB infection in children under 5 years: a systematic review with meta-analysis

S. Debulpaep,[1,2] H. Van Overschelde,[1,2] L. Pieters,[2] A. Soler-Garcia,[3] A. Noguera-Julian,[3,4,5] T. Goetghebuer,[6] P. Schelstraete,[2,7] M. Hainaut,[6] N.S. Pauwels[8]

[1]Department of Pediatrics, Ghent University Hospital, Ghent, Belgium; [2]Department of Internal Medicine and Pediatrics, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium; [3]Malalties Infeccioses i Resposta Inflamatòria Sistèmica en Pediatria, Servei d'Infectologia, Institut de Recerca Sant Joan de Déu, Barcelona, Spain; [4]Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain; [5]Departament de Cirurgia i Especialitats Medicoquirúrgiques, Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona, Barcelona, Spain; [6]Department of Pediatrics, Centre Hospitalier Universitaire Saint-Pierre, Université Libre de Bruxelles, Brussels, Belgium; [7]Department of Pediatric Pulmonology and Infectious Diseases, Ghent University Hospital, Ghent, Belgium; [8]Knowledge Centre for Health Ghent, Ghent University and Ghent University Hospital, Ghent, Belgium.

SUMMARY

BACKGROUND: Interferon-gamma release assays (IGRAs) have largely replaced the tuberculin skin test (TST) for diagnosing TB infection (TBI) in low-TB-burden countries, except in young children. This review assesses the diagnostic accuracy of IGRAs in children under 5 years old, using TST as a reference, for detecting TBI.

METHODS: A systematic review was conducted in MEDLINE, Embase, Web of Science, Scopus, and LILACS to identify studies published up to November 2024 that performed simultaneous IGRA and TST for TBI in children under five in low-TB-prevalence settings. Studies in mid- or high-burden countries involving immunocompromised children, TB disease, or non-TB mycobacteria, as well as studies lacking paired IGRA and TST, were excluded. A meta-analysis quantified the sensitivity and specificity of IGRA and explored heterogeneity.

RESULTS: Thirty-one reports met the inclusion criteria, providing 5,679 paired results. Compared to the TST, pooled sensitivity for IGRAs was 87.3% (95% confidence interval: 73.0–94.1), while specificity reached 98.3% (95% confidence interval: 95.7–99.2).

CONCLUSION: These findings support the use of IGRA for diagnosing TBI in young children in low-prevalence settings. We recommend a cautious approach in high-risk children, including the combination of both tests.

KEY WORDS: tuberculosis; low-TB-burden settings; paediatric; contact investigation; tuberculin skin test; TBI

About a quarter of the world's population is estimated to be infected with *Mycobacterium tuberculosis*, the immunological state of TB infection (TBI) without TB disease.[1,2] It is estimated that 5%–10% of individuals will develop TB disease over their lifetime, with half occurring in the first year after infection.[3] The risk of developing TB after infection is age-dependent[4]: infants (0–1 year) face the highest risk, with up to 50% developing TB, whereas this risk declines to 20% in children aged 1–5 years and 2% in children aged 5–10 years.[5,6] In 2023, an estimated 10.8 million people worldwide had TB; children accounted for 12% of all cases.[2] Due to high susceptibility, the WHO advises TB preventive treatment for children under the age of 5 years with contact exposure to contagious pulmonary TB, even if initial screening is negative, to lower the risk of progression to TB.[7] Detecting TBI in young children is challenging due to immature immune responses, delayed diagnosis and treatment, and increased risk of progression to severe TB.[8,9] In low-TB-burden settings, 22 children under 5 years of age with positive screening results need to be treated to prevent one case of TB.[10]

TBI diagnosis relies on immunodiagnostic testing: the in vivo tuberculin skin test (TST) and the in vitro interferon-gamma release assays (IGRAs). TST has poor specificity in the context of prior vaccination with bacille Calmette-Guérin (BCG) or non-TB mycobacteria (NTM) exposure, and low sensitivity in young, immunosuppressed children and individuals with severe TB.[11] IGRAs are more specific, unaffected by BCG vaccination and most NTM, but show weaker interferon-gamma responses in young children, increasing the risk of false-negative and indeterminate results.[12–14] TST interpretation depends on the epidemiological context and the selected cut-off.[15] A cut-off of ≥10 mm may miss individuals with a weaker immune response, such as children under five.[4] Lowering the cut-off to ≥5 mm increases sensitivity

SD and HVO contributed equally.

Correspondence to: Sara Debulpaep, Department of Pediatrics, Ghent University Hospital, Corneel Heymanslaan 10, 9000 Ghent, Belgium. E-mail: sara.debulpaep@uzgent.be

Article submitted 20 June 2025. Final version accepted 9 August 2025.

but reduces specificity, leading to more false-positive results. Region-specific guidelines reflect the prevalence of TB,[16] NTM exposure, and public health policies. In low-TB-burden countries, targeted screening helps identify and treat TBI in high-risk groups, reducing their risk of progression to TB. TST's positive predictive value (PPV) is lower in low-TB-burden countries than in high-burden regions due to the lower incidence of TB. Still, the negative predictive value (NPV) is higher.[17]

Most guidelines do not recommend using IGRAs alone for children under 5 years, with TST remaining the primary screening method.[18–21] However, since 2021, the American Academy of Pediatrics has recommended using IGRAs instead of TSTs for children older than 2 years.[14,22] A recent systematic review by Volkman et al.[23] evaluated QuantiFERON-TB Gold In-Tube (QFT-GIT) in children under five in various TB-burden settings, and concluded that this IGRA cannot be used to exclude TBI in this age group. We have focussed on low-TB-burden settings to minimise exposure variability, include all publications regardless of the IGRA used, and employ a correction model to account for diagnostic uncertainty. Our review aims to evaluate the accuracy of using IGRA in children under five in low-TB-burden settings to detect an *M. tuberculosis*–specific immunological response in the absence of symptoms or signs suggestive of TB, hereafter called *M. tuberculosis* infection.

METHODS

This review is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Diagnostic Test Accuracy (PRISMA-DTA) guidelines, and its protocol is registered and remains available on PROSPERO (CRD42024554531).[24] A comprehensive search was conducted in February 2024 and updated in November 2024 across five databases: MEDLINE (via PubMed), Embase (via Embase.com), Web of Science, Scopus, and LILACS. Additional searches on ClinicalTrials.gov, the WHO International Clinical Trials Registry Platform (ICTRP), and the International Standard Randomised Controlled Trials Number (ISRCTN) identified ongoing trials. All search strategies are outlined in Supplementary Data. Eligible studies compare IGRAs and TSTs in children under 5 years old in low-TB-burden countries (incidence $\leq$10 per 100,000 population) and are written in English, Dutch, French, or Spanish. Included IGRA tests are QFT-GIT, QuantiFERON-TB Plus (QFT-Plus) (QIAGEN), and T-SPOT.TB (Oxford Immunotec, now Revvity). We excluded studies from mid- and high-burden settings (incidence >10/100,000), those involving immunocompromised children, those focusing solely on TB disease or NTM infections, and those without paired IGRA and TST performed within 14 days.

Two reviewers (SD and HVO) independently screened all records for titles and abstracts and assessed the full texts using DistillerSR. Disagreements were resolved through discussion or by a third reviewer (NSP). Data extracted by two reviewers (SD, HVO, or LP) included the following: country, TB incidence, sample size, age, BCG vaccination rate, reason for investigation, and tests used. Data were extracted from published articles or, when unpublished, requested from study authors. Two reviewers (SD and HVO) independently assessed the risk of bias using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-II) tool.

Statistical analysis

Data analysis, including a meta-analysis, is performed using R statistical software (version 4.4.1) and WinBugs (version 1.4.3). Sensitivity and specificity were calculated for IGRA per study, reporting true positives, false positives, true negatives, and false negatives, assuming the reported TST results accurately reflected participants' true infection status. The kappa coefficient assessed the agreement between the TST (reference test) and the IGRAs (index test). Primary outcomes were sensitivity, specificity, PPV and NPV, agreement rate, and the kappa coefficient for IGRA, using TST as the reference in a 2 $\times$ 2 table. We employed a Bayesian bivariate meta-regression model, inspired by the bivariate binomial model described by Reitsma et al.[25] This model simultaneously estimates sensitivity and specificity while accounting for between-study heterogeneity. Confidence intervals (CIs) are derived by constructing a range around the point estimate, with the margin of error calculated as the product of the standard error and a critical value from the normal distribution. While CIs indicate the precision of the pooled estimates, prediction intervals (PIs) are calculated to reflect the expected variability in test performance across future settings or individual studies. A wide PI, particularly in terms of sensitivity, indicates substantial heterogeneity across studies. We apply two sets of assumptions:

1) Assuming TST is a perfect reference standard for TBI.
2) Assuming TST is an imperfect reference standard, with varying sensitivity and specificity. In this approach, for TST sensitivity, we applied correction factors based on published sensitivity estimates stratified by age group and TST induration cut-off for TB, as reported by Buonsenso et al.[26] For TST specificity, the estimates were derived from meta-analysis of data by Hamada et al.[17] These stratified published values were used to adjust study-level results, aiming better to reflect IGRA performance relative to an imperfect reference standard.

The meta-analytic model was also fitted to three subgroups: studies with low BCG coverage, high-quality

studies evaluated using QUADAS-II, and studies that used only QFT-GIT as IGRA. Ethical approval was not required as this study is a systematic review of previously published literature.

Ethical statement

Ethical approval was not required for this study as it is a systematic review of previously published literature and did not involve the collection of new data from human participants.

RESULTS

The PRISMA-DTA flow chart (Figure 1) illustrates the study selection. After removing duplicates using EndNote, 4,535 studies were screened based on their titles and abstracts. Of these, 575 reports underwent full-text review, with 31 studies meeting the inclusion criteria.[11–13,28–55]

Table 1 summarises the characteristics of the 31 included studies, published between 2008 and 2023 across 13 countries, which included data from

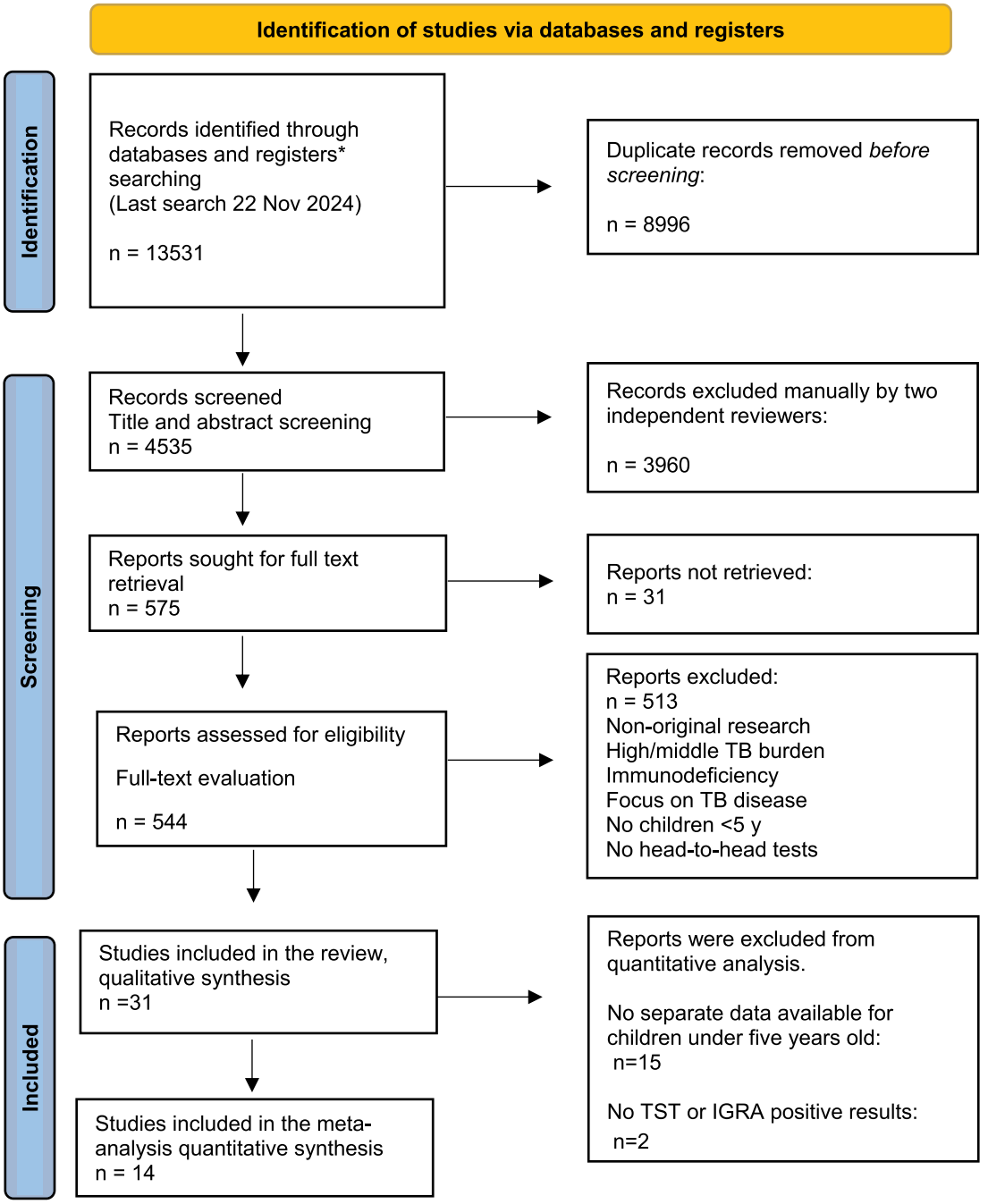

Figure 1. PRISMA-DTA 2020 flow diagram for systematic reviews. *MEDLINE (via PubMed), Embase (via Embase.com), Web of Science, Scopus, and LILACS. Additional searches on ClinicalTrials.gov, the World Health Organization International Clinical Trials Registry Platform (ICTRP), and the International Standard Randomised Controlled Trials Number (ISRCTN) identified ongoing trials. (Search strategies are added in Supplementary Data Appendix SA). Source: Page et al.[27]

Table 1. Characteristics of studies performing simultaneous tuberculin skin test and interferon-gamma release assay in children under 5 years of age.

Authors	Year	Country	TB incidence	n total	n < 5 years	Age (range)	BCG (%)	IGRA	TST cut-off	Reason for investigation
Ahmed et al.[28]	2020	USA	2.6	3,593	900	Median 8.8 years (IQR 5–11.8)	65.6	QFT-GIT, T-SPOT.TB	5–10 mm	TB contact investigation
Amanatidou et al.[29]	2015	Greece	2.2	161	63	Mean 6.3 years (SD 1.74)	24.8	QFT-GIT	5–10–15 mm	Clinical suspicion, TB contact investigation, recent immigration
Bergamini et al.[13]	2009	Italy	4.6	496	104	Mean 11.1 years (SD 5.7)	38.9	QFT-GIT, T-SPOT.TB, QFT-Plus	5–10 mm	Clinical suspicion, TB contact investigation, recent immigration
Basu Roy et al.[30]	2012	UK	7.6	1,128	268	Mean 8.2 years (SD 4.4)	60.6	QFT-GIT, T-SPOT.TB	10 mm	TB contact investigation, recent immigration
Bianchi et al.[11]	2009	Italy	4.6	336	152	Median 4.5 years (IQR 2.6–6.8)	54.5	QFT-GIT	5–10 mm	Clinical suspicion, TB contact investigation, recent immigration
Bua et al.[31]	2013	Italy	4.6	105	50	Range 3 months to 15 years	1.9	QFT-GIT	5 mm	TB contact investigation
Chiappini et al.[32]	2019	Italy	4,6	4,631	1,795	Median 5.7 years (IQR 5.6–5.8)	51.4	QFT-GIT	5–10–15 mm	Clinical suspicion, TB contact investigation, recent immigration
Chiappini et al.[33]	2018	Italy	4.6	762	Not stated	Mean 3, 6 years (IQR 3.5–5.7)	53.9	QFT-GIT	10 mm	Recent immigration
Chiappini et al.[34]	2014	Italy	4.6	338	133	Median 5.5 years (IQR 2.7–9.5)	33.1	QFT-GIT, T-SPOT.TB	Not stated	Clinical suspicion, TB contact investigation, recent immigration
Connell et al.[35]	2008	Australia	5.6	96	16	Mean 8.9 years (range 0.5–19)	52.0	QFT-GIT, T-SPOT.TB	5–10 mm	Clinical suspicion, TB contact investigation, recent immigration
Critselis et al.[12]	2012	Greece	2.2	761	198	Mean 7.84 years (SD 4.68)	45.2	QFT-GIT	5–10–15 mm	TB contact investigation, recent immigration
Debulpaep et al.[36]	2019	Belgium	7.8	60	60	Median 19.5 months (range 0.5–59)	18.3	QFT-GIT	5–10 mm	Clinical suspicion, TB contact investigation, recent immigration
Diel et al.[37]	2008	Germany	5.1	954	24	Mean 29.02 years (SD 11.8)	51.9	QFT-GIT	5 mm	TB contact investigation
Elliot et al.[38]	2018	Australia	5.6	212	68	Not stated	91.0	QFT-GIT	10 mm	Recent immigration
Erkens et al.[39]	2014	Netherlands	4.1	3,789	28	Not stated	40.2	QFT-GIT, T-SPOT.TB	5 mm	Clinical suspicion, TB contact investigation, recent immigration
Garazzino et al.[40]	2014	Italy	4.6	823	823	Median 13.5 months (IQR 8.4–18.9)	26.5	QFT-GIT	5–10 mm	Clinical suspicion, TB contact investigation, recent immigration
Lighter et al.[41]	2009	USA	2.6	207	67	Mean 9 years (SD 5.7)	35.7	QFT-GIT	10 mm	Clinical suspicion, TB contact investigation, recent immigration
Mastrolia et al.[42]	2019	Italy	4.6	1,775	1,020	Median 5.8 years (IQR 3.3–8.2)	76.0	QFT-GIT	10 mm	Recent immigration
Mendez et al.[43]	2011	Spain	6.9	459	194	Mean 4.73 years (SD 3.68)	46.4	QFT-GIT	5–10 mm	Clinical suspicion, TB contact investigation, recent immigration
Mulder et al.[44]	2019	Netherlands	4.1	49,990	8,516	Not stated	24.4	QFT-GIT, QFT-plus, ELISpot	5 mm	TB contact investigation, recent immigration
Pasqualini et al.[45]	2023	France	7.2	261	261	Not stated	67.0	Not stated	Not stated	TB contact investigation
Pavic et al.[46]	2015	Croatia	2.7	171	171	Mean 29 months (SD 16)	98.8	QFT-GIT	10 mm	TB contact investigation
Pop et al.[47]	2022	Switzerland	4.6	72	72	Not stated	Not stated	T-SPOT.TB	Not stated	TB contact investigation
Porter et al.[48]	2023	UK	7.6	241	Not stated	Median 9.1 years (IQR 3.3–12.6)	Not stated	Not stated	5–15 mm	TB contact investigation
Seddon et al.[49]	2016	UK	7.6	422	156	Median 69 months (IQR: 32–113)	71.1	QFT-GIT, T-SPOT.TB	5–10–15 mm	TB contact investigation
Soler-Garcia et al.[50]	2022	Spain	6.9	1,726	368	Median 8.4 years	17.6	QFT-plus	5–10 mm	Clinical suspicion, TB contact investigation, recent immigration
Sollai et al.[51]	2017	Italy	4.6	1,355	623	Median 5.3 years (IQR 3.0–7.9)	81.8	QFT-GIT	10 mm	Recent immigration adoption
Spicer et al.[52]	2015	USA	2.6	109	89	Median 22 months (range 4–193)	72.5	T-SPOT.TB	5–10 mm	Recent immigration
Stout et al.[53]	2018	USA	2.6	464	464	Median 3.1 years (IQR 2.2–4.0)	Not stated	QFT-GIT, T-SPOT.TB	5–10 mm	TB contact investigation, recent immigration
Thomas et al.[54]	2011	UK	7.6	283	69	Mean 5.29 years (SD 4.05)	71.7	QFT-GIT	5–15 mm	TB contact investigation
Velasco et al.[55]	2018	Spain	5.1	383	383	Median 29 months (IQR 17; 41)	26.1	QFT-GIT	5–10 mm	Clinical suspicion, TB contact investigation, recent immigration

BCG = bacille Calmette-Guérin vaccination; IGRA = interferon-gamma release assay; IQR = interquartile range; n = number; TST = tuberculin skin test; QFT-GIT = QuantiFERON Gold In-Tube; QFT-Plus = QuantiFERON Gold Plus; SD = standard deviation; TB Incidence: n per 100,000 population.

Table 2. Contingency table: tuberculin skin test and interferon-gamma release assay in children under 5 years of age.

Authors	Year	IGRA	TST cut-off	Indeterminate results	BCG coverage (%)	Quadas-II	Total n	TP n	FP n	FN n	TN n	Concordance	κ	Reason for investigation
Ahmed et al.[28]	2020	Combination	5–10 mm	Excluded (1.3%)	66%	Low risk of bias	898	32	7	198	661	77%	0.18	Contact investigation
Basu Roy et al.[30]	2012	QFT-GIT	10 mm	Not stated	62%	High risk of bias	257	47	75	12	123	66%	0.30	Mixed screening
Basu Roy et al.[30]	2012	T-SPOT.*TB*	10 mm	Not stated	62%	High risk of bias	260	22	100	2	135	60%	0.17	Mixed screening
Bianchi et al.[11]	2009	QFT-GIT	5–10 mm	Excluded (0.2%)	52%	Low risk of bias	336	36	24	45	229	79%	0.38	Mixed screening
Debulpaep et al.[36]	2019	QFT-GIT	5–10 mm	Excluded (3%)	18%	Low risk of bias	97	8	3	6	80	91%	0.59	Mixed screening
Erkens et al.[39]	2017	Combination	5 mm	Not stated	40%	High risk of bias	13	2	0	8	3	38%	0.10	Mixed screening
Garazzino et al.[40]	2014	QFT-GIT	5–10 mm	Excluded (4.2%)	27%	High risk of bias	616	96	25	29	466	91%	0.73	Mixed screening
Lighter et al.[41]	2009	QFT-GIT	10 mm	Excluded (1.4%)	36%	Low risk of bias	32	0	1	11	20	63%	−0.06	Mixed screening
Mastrolia et al.[42]	2019	QFT-GIT	10 mm	Excluded (0.4%)	76%	Low risk of bias	1,020	53	9	84	874	91%	0.49	Mixed screening
Mulder et al.[44]	2019	Combination	5 mm	Not stated	24%	High risk of bias	133	5	0	125	3	6%	0.00	Mixed screening
Pavic et al.[46]	2015	QFT-GIT	10 mm	Excluded (1.1%)	99%	Low risk of bias	171	18	8	13	132	88%	0.56	Contact investigation
Seddon et al.[49]	2016	Combination	5–10–15 mm	Included as negative	71%	Medium risk of bias	181	20	5	33	123	79%	0.40	Contact investigation
Soler-Garcia et al.[50]	2022	QFT-Plus	5–10 mm	Excluded (3.1%)	18%	Medium risk of bias	368	66	8	17	277	93%	0.80	Mixed screening
Stout et al.[53]	2018	QFT-GIT	5–10 mm	Not stated	Not stated	High risk of bias	464	15	2	116	331	75%	0.15	Mixed screening
Stout et al.[53]	2018	T-SPOT.*TB*	5–10 mm	Not stated	Not stated	High risk of bias	464	10	1	121	332	74%	0.10	Mixed screening
Velasco et al.[55]	2018	QFT-GIT	5–10 mm	Excluded (3.8%)	27%	Medium risk of bias	369	56	0	62	251	83%	0.55	Mixed screening

BCG = bacille Calmette-Guérin vaccination; IGRA = interferon-gamma release assay; IQR = interquartile range; κ = kappa; n = number; TST = tuberculin skin test; QFT-GIT = QuantiFERON Gold In-Tube; QFT-Plus = QuantiFERON Gold Plus; TP = true positive: TST+/IGRA+; FP = false positive: TST−/IGRA+ results; FN = false negative: TST+/IGRA−; TN = true negative: TST−/IGRA−.

76,163 children under five in low-TB-burden settings. Most studies (71%) were prospective, and 93% used QFT-GIT as IGRA. Seven studies (22%) used multiple IGRAs.[13,28,34,35,39,44,49] TST cut-offs varied (5, 10, 15 mm), reflecting public health policies. TST-IGRA was performed for mixed reasons: most studies (81%) included contact investigations for exposed children, 58% included recent immigration from high-burden areas, and 41% included investigations after clinical suspicion.

Risk of bias was assessed using the QUADAS-II tool (Supplementary Data Appendix SB). Overall, study quality was moderate to good. We considered using TST as a reference test, given its potential for positivity due to factors other than *M. tuberculosis* infection (e.g., TB disease, NTM infection, and BCG vaccination), and we judged it to be at high risk of bias across all studies.

Sixteen studies providing paired TST-IGRA results were screened.[11,28,30,36,39–42,44,46,47,49,50,52,53,55] Two studies, Pop et al.[47] and Spicer et al.,[52] lacked any TST-positive or IGRA-positive cases and were excluded from further analysis, as sensitivity and specificity could not be calculated from their data. The remaining 14 studies were included in the meta-analysis to construct a 2 × 2 contingency matrix, as shown in Table 2.[11,28,30,36,39–42,44,46,49,50,53,55] Two studies reported separate results for QFT-GIT and T-SPOT.TB and were therefore analysed individually.[30,53] All data were aggregated; individual participant data were unavailable. Concordance between TST and IGRA varied across studies, with agreement rates ranging from 60% to 93%; two studies reported kappa coefficients greater than 0.7.[40,50] Across all included studies with calculable kappa values (Table 2), agreement between TST and IGRA ranged from poor to substantial ($\kappa = -0.06$ to 0.80). Most studies have tested children for mixed screening indications, often including, but not limited to, contact investigation. Studies conducted exclusively in the context of contact investigation (n = 3)[28,46,49] showed higher agreement (mean agreement 81%; κ 0.38) than those with mixed screening indications (n = 11)[11,30,36,39–42,44,50,53,55] (mean agreement 68%; κ 0.31). The Mann–Whitney *U* test comparing kappa values between the two groups (contact investigation, n = 3, vs. mixed screening, n = 11) was not statistically significant ($P = 0.54$).

Finally, 14 studies and 5,679 paired TST-IGRA results were included in the meta-analysis (Figure 2). TST sensitivity was assigned using Buonsenso et al.[26] estimates based on age and used TST cut-off: 58% for studies involving younger children (0–2 years[40,41]), 90% for studies involving children aged 0–5 years with TST cut-off ≥5 mm,[39,44] 79% for studies involving children aged 0–5 years that did apply a combination of 5 and 10 mm cut-off, depending on the underlying indication for screening,[28,30,36,42,46,49,50,55] and TST specificity was assumed to be 76% across all studies following Hamada et al.[17] values. For Stout et al.,[53] we used the TST sensitivity and specificity estimates derived from the latent class analysis, as reported in the original publication. We then applied Bayesian Latent Class Modelling to refine these results further and calculate IGRA sensitivity and specificity.

This analysis highlighted substantial variability in IGRAs' sensitivity estimates across studies. After applying the Bayesian correction model, the pooled sensitivity and specificity for IGRAs were estimated at 87.3% sensitivity (95% CI: 73.0–94.1) with a PI of 29.9%–99.1% and 98.3% specificity (95% CI: 95.7–99.2), with a PI of 90.3%–99.7%, 1111,28,30,36,40,41,42,46,49,50,53,55 as shown in Table 3. Therefore, we implemented a stepwise approach to address this variability and refine our analysis. We conducted subgroup analyses. In the first subgroup, we included studies that solely used QFT-GIT as IGRA[11,30,36,40–42,46,53,55]; the second subgroup comprised studies with BCG coverage below 30% [36,40,50,55] and the third subgroup focused on studies with a high QUADAS-II score.[11,28,36,41,42,46]

1. QFT-GIT subgroup (n = 9): Sensitivity remained similar at 87.9% (95% CI: 63.6–95.6), but the PI was even broader (18.5%–99.5%), suggesting continued variability despite focusing on a single IGRA test.
2. Studies with a BCG coverage below 30% subgroup (n = 4): Sensitivity increased to 91.4% (95% CI: 37.3–95.8), but the PI widened significantly (0.6%–99.9%), likely due to the smaller number of included studies.
3. High-quality (QUADAS-II) subgroup (n = 6): Sensitivity was lower at 80.8% (95% CI: 34.4–94.9), but again, the PI remained large (3.9%–99.7%).

DISCUSSION

Diagnosing TBI in children under five remains complex due to immature immune responses and the limitations of indirect tests.[9] This review assessed the diagnostic accuracy of IGRA for the detection of *M. tuberculosis* infection compared to TST in low-TB-burden settings, synthesising data from 14 studies with 5,679 paired TST-IGRA results. IGRA specificity was high (98.3%), while sensitivity varied (87.3%), with a broad PI (29.9%–99.1%), reflecting heterogeneity across populations (e.g., IGRA type, BCG status, reason for testing, comorbidities, and TB exposure).

Subgroup analyses, restricted to QFT-GIT-only studies, studies with < 30% BCG-vaccinated participants, and high-quality studies, further widened the PI, particularly for sensitivity, indicating substantial heterogeneity between studies. This was also influenced by the smaller number of included studies,

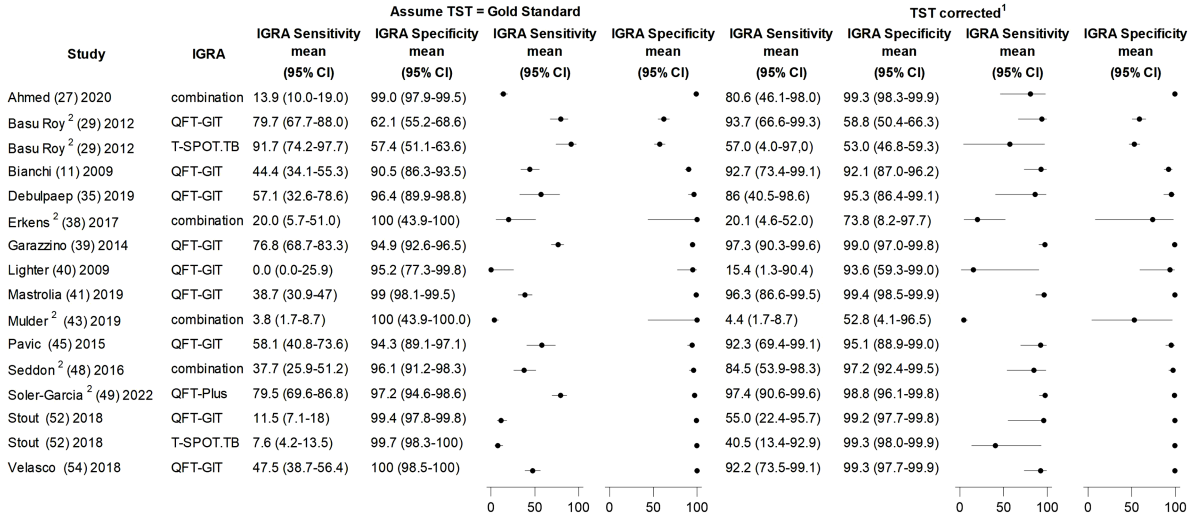

Figure 2. IGRA sensitivity and specificity results with confidence interval (CI) per study. [1]Correction factor based on published sensitivity by Buonsenso et al.[26] and specificity by Hamada et al.[17] [2]Calculated values based on non-peer-reviewed data retained after contacting the author. TST = tuberculin skin test; IGRA = interferon-gamma release assay; QFT-GIT = QuantiFERON Gold In-Tube; 95% CI = 95% confidence interval.

which increased statistical uncertainty. The <30% BCG subgroup, studies with less than 30% of the study population vaccinated with BCG, aimed to reduce the influence of BCG-related positive TST results. Without individual participant data on BCG vaccination, a population-level proxy was used. In this subgroup, the estimated IGRA sensitivity increased to 91.4%, likely reflecting a reduction in misclassification of true infection status due to BCG cross-reactivity, as indicated by positive TSTs. Conversely, high-quality studies showed lower sensitivity (80.8%), potentially reflecting more rigorous methodologies that better expose the limitations inherent to IGRAs, such as suboptimal sensitivity in young children and immunocompromised individuals, the inability to distinguish between *M. tuberculosis* infection and TB, and susceptibility to sample handling issues. IGRAs also rely on a single-time-point immune response, making them sensitive to fluctuating immune status and timing.

The systematic review by Volkman et al.[23] reported a lower IGRA sensitivity (45%) and specificity (96%), with moderate inter-study agreement (average kappa = 0.50, range 0.17–0.80). Compared to our findings, their analysis showed lower sensitivity, which may reflect the inclusion of settings with higher TB burden, potentially increasing the proportion of children with TB at testing. Differences in IGRA scope (QFT-GIT only), the reference standard, and statistical methods (latent class analysis) may further explain the differences. The comparison underscores the importance of study context in diagnostic accuracy reviews.

The key strengths of our review are its inclusion of all commercially available IGRAs and its use of a Bayesian correction model to account for TST limitations as a reference standard.[25] However, a fundamental limitation in evaluating diagnostic tests for *M. tuberculosis* infection remains the lack of a reference standard. Ideally, the reference standard would be the evolution to TB disease among untreated test-positive children; such data are understandably scarce in children under five, given their high risk of developing severe TB and therefore the recommendation to treat preventively. The study by Ahmed et al.,[28] which found no progression to TB among TST-positive, IGRA-negative BCG-vaccinated children, offers a

Table 3. Meta-analysis: sensitivity and specificity for IGRA using the correction factor.

	Sensitivity					Specificity				
		95% CI		95% PI			95% CI		95% PI	
	%	Low	High	Low	High	%	Low	High	Low	High
Total[A]	87.3	73	94.1	29.9	99.1	98.3	95.7	99.2	90.3	99.7
Subgroup QFT-GIT[B]	87.9	63.6	95.6	18.5	99.5	97.7	91.4	99.1	77.0	99.8
Subgroup low vaccination status[C]	91.4	37.3	95.8	0.6	99.9	97	79.1	99.2	44.0	99.9
Subgroup QUADAS-II high-quality[D]	80.8	34.4	94.9	3.9	99.7	96.8	82.5	99.1	53.6	99.9

QFT-GIT = QuantiFERON Gold In-Tube; CI = confidence interval; PI = prediction interval.
[A]Total: Ahmed et al.,[28] Bianchi et al.,[11] Basu Roy et al.,[30] Debulpaep et al.,[36] Garazzino et al.,[40] Lighter et al.,[41] Mastrolia et al.,[42] Pavic et al.,[46] Seddon et al.,[49] Soler-Garcia et al.,[50] Stout et al.,[53] and Velasco et al.[55] [B]Included studies in meta-analysis QFT-GIT: Bianchi et al.,[11] Debulpaep et al.,[36] Garazzino et al.,[40] Lighter et al.,[41] Mastrolia et al.,[42] Pavic et al.,[46] Stout et al.,[53] and Velasco et al.[55] [C]Included studies in meta-analysis low vaccination status <30%: Debulpaep et al.,[36] Garazzino et al.,[40] Soler-Garcia et al.,[50] and Velasco et al.[55] [D]The following studies were included in the meta-analysis with high QUADAS-II: Ahmed et al.,[28] Bianchi et al.,[11] Debulpaep et al.,[36] Lighter et al.,[41] Mastrolia et al.,[42] and Pavic et al.[46]

clinically relevant perspective. Discordance between TST and IGRA is particularly well-documented in young, BCG-vaccinated children.[28,56] While guidelines advise against adjusting for BCG status in this age group to preserve sensitivity, this compromises specificity. Because individual-level data were not available, we used study-level BCG coverage as a proxy. Other limitations include population-level heterogeneity and regional disparities in low-incidence countries, including poverty, overcrowding, and limited health care access.[57] In several studies, the inclusion of recently immigrated children with diverse TB exposure histories and BCG coverage may further influence test results.[11,30,36,39–42,44,50,53,55] Finally, although we aimed to reduce bias by including only studies with paired TST-IGRA results, differences in study design may have impacted comparability. Our review highlights the need for more longitudinal data to guide the use of IGRA in children under five. Prospective studies or registries that track the development of TB in untreated children, particularly those who are TST-positive but IGRA-negative, as shown in the study by Ahmed et al.,[28] are crucial for improving risk stratification and supporting evidence-based decisions in this vulnerable age group.

CONCLUSION

This review demonstrates that IGRAs are sufficiently accurate to support their use in screening for *M. tuberculosis* infection among children under 5 years of age living in low-incidence countries. While the findings support the careful use of IGRAs in young children, particularly in screening contexts, they also highlight essential limitations. Notably, sensitivity remains variable across studies, and the risk of false-negative results, especially in children under 2 years or those with intense exposure, cannot be overlooked.

In high-risk children, we recommend a cautious diagnostic approach, including the combined use of TST and IGRA, to enhance diagnostic accuracy and ensure timely clinical decision-making. The wide PI observed highlights the importance of context in interpreting results. Young children should not be excluded from testing solely based on age; instead, they should be approached with diagnostic strategies tailored to their specific risk profile and clinical context. Our findings support the use of IGRA in young children, provided it is applied carefully and interpreted in conjunction with clinical judgement.

Acknowledgements

We thank J. Menten, biostatistician (Belgium), for conducting the meta-analysis. We also thank R.B. Roy (Queen Mary University of London, UK), Conny Erkens and Christiaan Mulder (KNCV Tuberculosis Foundation, The Hague, the Netherlands), J.A. Seddon (Stellenbosch University, Cape Town, South Africa, and Imperial College London, London, UK), and J.-P. Zellweger (TB Competence Center, Swiss Lung Association, Bern, Switzerland) for sharing unpublished data as collaborators of this publication. We thank M. Debulpaep (Vrije Universiteit Brussel, Belgium) for creating the forest plots. We gratefully acknowledge the support of the University Foundation of Belgium. This work was partially supported by a research grant from the Carlos III Institute of Health, Ministry of Economy and Competitiveness of Spain (PI22/00766 to ANJ), and by the Spanish Society of Pneumonology and Thoracic Surgery (169-2022 to ANJ).

Conflicts of interest: ANJ has received speaker honoraria from QIAGEN and is the principal investigator of a study sponsored by bioMérieux; all other authors have no conflicts of interest.

References

1. Cohen A, et al. The global prevalence of latent tuberculosis: a systematic review and meta-analysis. Eur Respir J. 2019;54(3): 1900655.
2. World Health Organization (WHO). Global tuberculosis report 2024. Geneva: WHO, 2024.
3. Ghosh S, Dronavalli M, Raman S. Tuberculosis infection in under-2-year-old refugees: should we be screening? A systematic review and meta-regression analysis. J Paediatr Child Health. 2020; 56(4):622-629.
4. Sloot R, et al. Risk of tuberculosis after recent exposure. A 10-year follow-up study of contacts in Amsterdam. Am J Respir Crit Care Med. 2014;190(9):1044-1052.
5. Diel R, et al. Interferon-gamma release assays for the diagnosis of latent Mycobacterium tuberculosis infection: a systematic review and meta-analysis. Eur Respir J. 2011;37(1):88-99.
6. Marais BJ, et al. Childhood pulmonary tuberculosis: old wisdom and new challenges. Am J Respir Crit Care Med. 2006;173(10): 1078-1090.
7. Lönnroth K, Raviglione M. The WHO's new end TB strategy in the post-2015 era of the sustainable development goals. Trans R Soc Trop Med Hyg. 2016;110(3):148-150.
8. Auguste P, et al. Identifying latent tuberculosis in high-risk populations: systematic review and meta-analysis of test accuracy. Int J Tuberc Lung Dis. 2019;23(11):1178-1190.
9. Vanden Driessche K, et al. Immune vulnerability of infants to tuberculosis. Clin Dev Immunol. 2013;2013:781320.
10. Martinez L, et al. Effectiveness of preventive treatment among different age groups and Mycobacterium tuberculosis infection status: a systematic review and individual-participant data meta-analysis of contact tracing studies. Lancet Respir Med. 2024; 12(8):633-641.
11. Bianchi L, et al. Interferon-gamma release assay improves the diagnosis of tuberculosis in children. Pediatr Infect Dis J. 2009; 28(6):510-514.
12. Critselis E, et al. The effect of age on whole blood interferon-gamma release assay response among children investigated for latent tuberculosis infection. J Pediatr. 2012;161(4):632-638.
13. Bergamini BM, et al. Performance of commercial blood tests for the diagnosis of latent tuberculosis infection in children and adolescents. Pediatrics. 2009;123(3):e419-e424.
14. Lewinsohn DM, et al. Official American thoracic society/infectious diseases society of America/centers for disease control and prevention clinical practice guidelines: diagnosis of tuberculosis in adults and children. Clin Infect Dis. 2017;64(2):111-115.
15. Kik SV, et al. Predictive value for progression to tuberculosis by IGRA and TST in immigrant contacts. Eur Respir J. 2010;35(6): 1346-1353.
16. World Health Organization Tuberculosis Research Office. Further studies of geographic variation in naturally acquired tuberculin sensitivity. Bull World Health Organ. 1955;12(1-2):63-83.
17. Hamada Y, et al. Predictive performance of interferon-gamma release assays and the tuberculin skin test for incident tuberculosis: an individual participant data meta-analysis. EClinicalMedicine. 2023;56:101815.
18. World Health Organization. WHO consolidated guidelines on tuberculosis: module 2: screening: systematic screening for tuberculosis disease. Geneva: WHO, 2021.

19. Turnbull L, Bell C, Child F. Tuberculosis (NICE clinical guideline 33). Arch Dis Child Educ Pract Ed. 2017;102(3):136-142.
20. Saukkonen JJ, et al. Updates on the treatment of drug-susceptible and drug-resistant tuberculosis: an official ATS/CDC/ERS/IDSA clinical practice guideline. PLoS Negl Trop Dis. 2025;211(1):15-33.
21. Migliori GB, et al. ERS/ECDC statement: European Union standards for tuberculosis care, 2017 update. Eur Respir J. 2018;51(5):1702678.
22. Nelson D, Starke JR. Committee on Infectious Diseases. Tuberculosis infection in children and adolescents: testing and treatment. Pediatrics. 2021;148(6).
23. Volkman T, et al. QuantiFERON Gold-In-Tube for the diagnosis of mycobacterial tuberculosis infection in children under 5 years of age: a systematic review and meta-analysis. PLoS One. 2024;19(1):e0295913.
24. McInnes MD, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. JAMA. 2018;319(4):388-396.
25. Reitsma JB, et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2025;58(10):982-990.
26. Buonsenso D, et al. Performance of QuantiFERON-TB gold plus assays in paediatric tuberculosis: a multicentre PTBNET study. Thorax. 2023;78(3):288-296.
27. Page MJ, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
28. Ahmed A, et al. Interferon-γ release assays in children <15 years of age. Pediatrics. 2020;145(1):e20191930.
29. Amanatidou V, et al. Interferon gamma inducible protein-10 in the diagnosis of paediatric tuberculosis infection in a low TB incidence country. Int J Tuberc Lung Dis. 2015;19(12):1463-1469.
30. Basu Roy R, et al. Identifying predictors of interferon-γ release assay results in pediatric latent tuberculosis: a protective role of bacillus Calmette-Guerin?: a pTB-NET collaborative study. Am J Respir Crit Care Med. 2012;186(4):378-384.
31. Bua A, et al. Tuberculin skin test and QuantiFERON in children. New Microbiol. 2013;36(2):153-156.
32. Chiappini E, et al. QuantiFERON-TB gold in-tube test performance in a large pediatric population investigated for suspected tuberculosis infection. Paediatr Respir Rev. 2019;32:36-47.
33. Chiappini E, et al. Italian multicentre study found infectious and vaccine-preventable diseases in children adopted from Africa and recommends prompt medical screening. Acta Paediatr. 2018;107(9):1518-1586.
34. Chiappini E, et al. Interferon-gamma release assay sensitivity in children younger than 5 years is insufficient to replace the use of tuberculin skin test in western countries. Pediatr Infect Dis J. 2014;33(12):1291-1293.
35. Connell TG, et al. A three-way comparison of tuberculin skin testing, QuantiFERON-TB gold and T-SPOT.TB in children. PLoS One. 2008;3(7):e2624.
36. Debulpaep S, et al. Contribution of QuantiFERON-TB gold-in-tube to the diagnosis of Mycobacterium tuberculosis infection in young children in a low TB prevalence country. Front Pediatr. 2019;7:291.
37. Diel R, et al. Predictive value of a whole blood IFN-gamma assay for the development of active tuberculosis disease after recent infection with Mycobacterium tuberculosis. Am J Respir Crit Care Med. 2008;177(10):1164-1170.
38. Elliot C, et al. Tuberculin skin test versus interferon-gamma release assay in refugee children: a retrospective cohort study. J Paediatr Child Health. 2018;54(8):834-839.
39. Erkens CG, et al. Added value of interferon-gamma release assays in screening for tuberculous infection in the Netherlands. Int J Tuberc Lung Dis. 2014;18(4):413-420.
40. Garazzino S, et al. Performance of interferon-γ release assay for the diagnosis of active or latent tuberculosis in children in the first 2 years of age: a multicenter study of the Italian Society of Pediatric Infectious Diseases. Pediatr Infect Dis J. 2014;33(9):e226-e231.
41. Lighter J, et al. Latent tuberculosis diagnosis in children by using the QuantiFERON-TB gold in-tube test. Pediatrics. 2009;123(1):30-37.
42. Mastrolia MV, et al. Utility of tuberculin skin test and IGRA for tuberculosis screening in internationally adopted children: retrospective analysis from a single center in Florence, Italy. Travel Med Infect Dis. 2019;28:64-67.
43. Méndez-Echevarría A, et al. Optimizing interpretation of the tuberculin test using an interferon-gamma release assay as a reference standard. Pediatr Infect Dis J. 2011;30(5):426-428.
44. Mulder C, et al. Tuberculin skin test reaction depends on type of purified protein derivative: implications for cut-off values. Int J Tuberc Lung Dis. 2019;23(12):1327-1334.
45. Pasqualini C, et al. Tuberculosis in 0-5-year-old children following TB contact investigations: a retrospective study in a low burden setting. Front Pediatr. 2023;11:1145191.
46. Pavic I, et al. Discordance between tuberculin skin test and interferon-gamma release assay in children younger than 5 Years who have been vaccinated with Bacillus calmette-guerin. Lab Med. 2015;46(3):200-206.
47. Pop R, et al. Low secondary attack rate after prolonged exposure to sputum smear positive miliary tuberculosis in a neonatal unit. Antimicrob Resist Infect Control. 2022;11(1):148.
48. Porter GJ, et al. Optimizing the threshold to treat children for latent tuberculosis infection: a North-East England experience. Pediatr Infect Dis J. 2023;42(3):e88-e89.
49. Seddon JA, et al. The impact of BCG vaccination on tuberculin skin test responses in children is age dependent: evidence to be considered when screening children for tuberculosis infection. Thorax. 2016;71(10):932-939.
50. Soler-Garcia A, et al. Performance of QuantiFERON-TB Gold Plus assays in children and adolescents at risk of tuberculosis: a cross-sectional multicentre study. Thorax. 2022;77(12):1193-1201.
51. Sollai S, et al. Infectious diseases prevalence, vaccination coverage, and diagnostic challenges in a population of internationally adopted children referred to a Tertiary Care Children's Hospital from 2009 to 2015. Med Baltim. 2017;96(12):e6300.
52. Spicer KB, et al. Tuberculin skin testing and T-SPOT. TB in internationally adopted children. Pediatr Infect Dis J. 2015;34(6):599-603.
53. Stout JE, et al. Evaluating latent tuberculosis infection diagnostics using latent class analysis. Thorax. 2018;73(11):1062-1070.
54. Thomas B, et al. Concordance between tuberculin skin test and interferon-γ assay and interferon-γ response to mitogen in pediatric tuberculosis contacts. Pediatr Pulmonol. 2011;46(12):1225-1232.
55. Velasco-Arnaiz E, et al. Performance of tuberculin skin tests and interferon-γ release assays in children younger than 5 years. Pediatr Infect Dis J. 2018;37(12):1235-1241.
56. Ho CS, et al. Comparison of three tests for latent tuberculosis infection in high-risk people in the USA: an observational cohort study. Lancet Infect Dis. 2022;22(1):85-96.
57. Newton SM, et al. Paediatric tuberculosis. Lancet Infect Dis. 2008;8(8):498-510.

SUMMARY

Q3

RESUME (Français): Les tests appelés Interferon-gamma release assays ont largement remplacé les tests cutanés à la tuberculine pour le diagnostic de l'infection avec *Mycobacterium tuberculosis* dans les pays à faible charge TB, à l'exception des jeunes enfants. Cette revue évalue la précision diagnostique des Interferon-gamma release assays chez les enfants de moins de cinq ans par rapport à l'intradermo-réaction à la tuberculine comme test de référence pour la détection de l'infection à *Mycobacterium tuberculosis*.

MÉTHODES: Une revue systématique a été réalisée dans les bases de données MEDLINE, Embase, Web of Science, Scopus et LILACS afin d'identifier les études publiées jusqu'en novembre 2024 ayant réalisé simultanément des Interferon-gamma release assays et des test cutané à la tuberculine pour le diagnostic de l'infection à *M. tuberculosis* chez les enfants de moins de cinq ans dans des contextes de faible prévalence de la tuberculose. Les études menées dans des pays à prévalence tuberculeuse moyenne ou élevée, celles portant sur des enfants immunodéprimés, atteints de tuberculose maladie ou d'infections à mycobactéries non tuberculeuses, ainsi que celles ne comportant pas de tests appariés entre Interferon-gamma release assays et des tests cutanés à la tuberculine ont été exclues. Une méta-analyse a permis d'estimer la sensibilité et la spécificité des Interferon-gamma release assays et d'explorer les sources d'hétérogénéité. Cette revue est enregistrée sur PROSPERO (CRD42024554531).

RÉSULTATS: Trente et une études ont été incluses, représentant un total de 5,679 résultats appariés. Par rapport à l'intradermo-réaction à la tuberculine, la sensibilité combinée des Interferon-gamma release assays était de 87.3 % (intervalle de confiance à 95%: 73.0–94.1) et la spécificité de 98.3 % (intervalle de confiance à 95%: 95.7–99.2).

CONCLUSION: Ces résultats soutiennent l'utilisation des Interferon-gamma release assays pour le diagnostic de l'infection *M. tuberculosis* chez les jeunes enfants dans des contextes de faible prévalence. Nous recommandons toutefois une approche diagnostique prudente chez les enfants à haut risque, incluant l'utilisation combinée de l'intradermo-réaction à la tuberculine et des Interferon-gamma release assays.

QUERIES AND CORRECTIONS

AUTHOR: PLEASE ANSWER ALL QUERIES THAT APPEAR BELOW.

Q1: Please confirm that the spelling and format of all author names is correct.

Q2: Please check the affiliations are displayed correctly.

Q3: As per style, Introduction heading is not allowed, so we have removed the 'RESUME (Français)'. Please check.

Q4: References 27–57 were out of numerical order, so they have been renumbered in the text and reference list. Could you please confirm whether the renumbering has been done correctly or amend as needed?