

Use of T Cell–Based Diagnosis of Tuberculosis Infection to Optimize Interpretation of Tuberculin Skin Testing for Child Tuberculosis Contacts

Mustafa Bakir,^{1,a} Davinder P. S. Dosanjh,^{6,a} Jonathan J. Deeks,^{7,a} Ahmet Soysal,¹ Kerry A. Millington,^{5,8} Serpil Efe,¹ Yasemin Aslan,¹ Dilek Polat,² Nihat Kodalli,³ Aysegul Yagci,⁴ Isil Barlan,¹ Nerin Bahceciler,¹ Emel E. Demiralp,⁵ and Ajit Lalvani^{6,8}

¹Department of Paediatrics, Marmara University School of Medicine, ²Umraniye Tuberculosis Clinic of the Turkish Ministry of Health, and Departments of ³Radiology, ⁴Microbiology, and ⁵Immunology, Marmara University School of Medicine, Istanbul, Turkey; and ⁶Nuffield Department of Medicine, University of Oxford, John Radcliffe Hospital, Oxford, ⁷Unit of Public Health, Epidemiology, and Biostatistics at the University of Birmingham, Birmingham, and ⁸Tuberculosis Research Unit, Department of Respiratory Medicine, National Heart and Lung Institute, Imperial College London, London, United Kingdom

Background. Treatment of recent tuberculosis infection in children aged <2 years is essential, because of high risk of progression to disease, but diagnosis is hindered by the inaccuracy of the tuberculin skin test (TST). More-accurate T cell–based tests of infection could enhance diagnosis by optimizing interpretation of the TST results.

Methods. A total of 979 child tuberculosis contacts in Istanbul underwent the TST and enzyme-linked immunospot assay. Using enzyme-linked immunospot test results as a reference standard, we assessed the effect of age and bacille Calmette-Guérin (BCG) vaccination on the sensitivity and specificity of the TST, and we computed the optimal TST cutoff points, using receiver operating characteristic curves.

Results. With a TST cutoff point of ≥ 10 mm, the sensitivity of the TST was 66% for children aged <2 years, which was lower than that for older children ($P = .006$). Specificity was 75% for BCG-vaccinated children, compared with 92% for unvaccinated children ($P = .001$). Optimal cutoff points improved TST specificity for children with 1 BCG scar, with little loss of sensitivity. Despite the use of optimal cutoff points, TST sensitivity remained <70% for children aged <2 years, specificity remained <87% for BCG-vaccinated children aged ≥ 2 years, and overall accuracy was low for children with >1 BCG scar.

Conclusions. Negative results of the TST cannot exclude tuberculosis infection for child tuberculosis contacts aged <2 years, which supports the use of preventive therapy regardless of the TST results for this age group. In children aged ≥ 2 years, the accuracy of the TST can be improved by adjustment of cutoff points for BCG-vaccinated children but remains poor for children with >1 BCG scar. This methodology can define optimal TST cutoff points for diagnosis of tuberculosis infection tailored to target populations.

Children with *Mycobacterium tuberculosis* infection are at high risk of developing active tuberculosis [1, 2]. Infected children aged <2 years have a $\geq 20\%$ risk of progression to active disease, which increases to 50% for infants [2]. The risk decreases among children aged

≥ 2 years and is $\sim 5\%$ for children aged 2–5 years [2]. Given that isoniazid therapy prevents progression to tuberculosis [3], rapid and accurate diagnosis of childhood *M. tuberculosis* infection is a global public health priority.

However, routine diagnosis of *M. tuberculosis* infection is based on the tuberculin skin test (TST), which handicaps management of childhood tuberculosis infection. Prior bacille Calmette-Guérin (BCG) vaccination may affect the specificity of the TST [4], and some countries adjust TST cutoff points for vaccinated children. However, there is no consensus as to whether or how much the TST cutoff point should be increased [5, 6].

The sensitivity of the TST for young children is un-

Received 12 August 2008; accepted 20 October 2008; electronically published 5 January 2009.

^a M.B., D.P.S.D., and J.J.D. contributed equally to this article.

Reprints or correspondence: Prof. Ajit Lalvani, Tuberculosis Research Unit, Dept. of Respiratory Medicine, National Heart and Lung Institute, St. Mary's Campus, Imperial College London, Norfolk Pl., London W2 1PG, United Kingdom (a.lalvani@imperial.ac.uk).

Clinical Infectious Diseases 2009;48:302–12

© 2009 by the Infectious Diseases Society of America. All rights reserved.

1058-4838/2009/4803-0006\$15.00

DOI: 10.1086/595847

known; therefore, guidelines for the management of child tuberculosis contacts vary widely. Some national tuberculosis control programs recommend preventive therapy with isoniazid for young children, on the basis of TST results [5, 7], whereas others recommend it for all child tuberculosis contacts younger than a certain age, regardless of TST results [6, 8–11].

T cell–based IFN- γ release assays (IGRAs) represent a significant advancement from the TST [12–15], and, in this study, the IFN- γ enzyme-linked immunospot assay (ELISPOT) was used. Although further research is required, much published evidence indicates that, for adults, ELISPOT is more specific and more sensitive than is the TST [16–28]. Although there are fewer studies involving children, the available evidence shows that ELISPOT is more sensitive than is the TST for children with active tuberculosis [29–31]. In addition, although there is no reference standard test for latent tuberculosis infection, ELISPOT appears to be more sensitive than the TST for contact investigations among children and infants [17, 30, 32–34]. Thus, available evidence suggests that ELISPOT is more specific and more sensitive than the TST for detection of latent tuberculosis infection in children.

However, IGRAs are not yet suitable for all populations, because they are expensive and require a blood sample and laboratory equipment. Therefore, most children may not benefit from this important medical advance. We reasoned that IGRAs might be used to inform and improve the interpretation and diagnostic accuracy of the TST. In the absence of a reference standard, we used ELISPOT as a surrogate reference standard for testing for latent tuberculosis infection, to optimize the use of the TST for children with recent tuberculosis exposure. ELISPOT was applied in parallel with the TST for 979 child household contacts in Istanbul, Turkey, which has an intermediate prevalence of tuberculosis (40 cases per 100,000 persons) [35], a very low prevalence of HIV infection among children [36], and a policy of universal childhood BCG vaccination. We classified ELISPOT-positive children as infected and ELISPOT-negative children as uninfected and then assessed the effects of young age and BCG vaccination on the sensitivity and specificity of the TST.

MATERIALS AND METHODS

Study participants. All adults who received a diagnosis of sputum smear–positive pulmonary tuberculosis at the 7 government-funded tuberculosis clinics in east Istanbul from October 2002 through May 2004 and who had children living in the household were invited to participate in the study, as described elsewhere [30]. Ethics approval was granted by the Institutional Review Board of Marmara University School of Medicine, Istanbul; The Turkish Ministry of Health, Ankara; and the World Health Organization Steering Committee on Research Involving Human Subjects, Geneva.

The Turkish Ministry of Health guidelines for BCG vaccination are as follows: all children are vaccinated intradermally with BCG Pasteur 1173-P2 (Serum Institute of India) at age 2–3 months, and a booster vaccination is administered during the first year of primary school, at age 6–7 years. The BCG vaccination coverage rate among Turkish children was 79% in 2004 [37].

Clinical evaluation. A total of 1024 child contacts of the 443 index patients with sputum smear–positive pulmonary tuberculosis were enrolled at the Paediatric Infectious Diseases Clinic, Marmara University School of Medicine, where medical

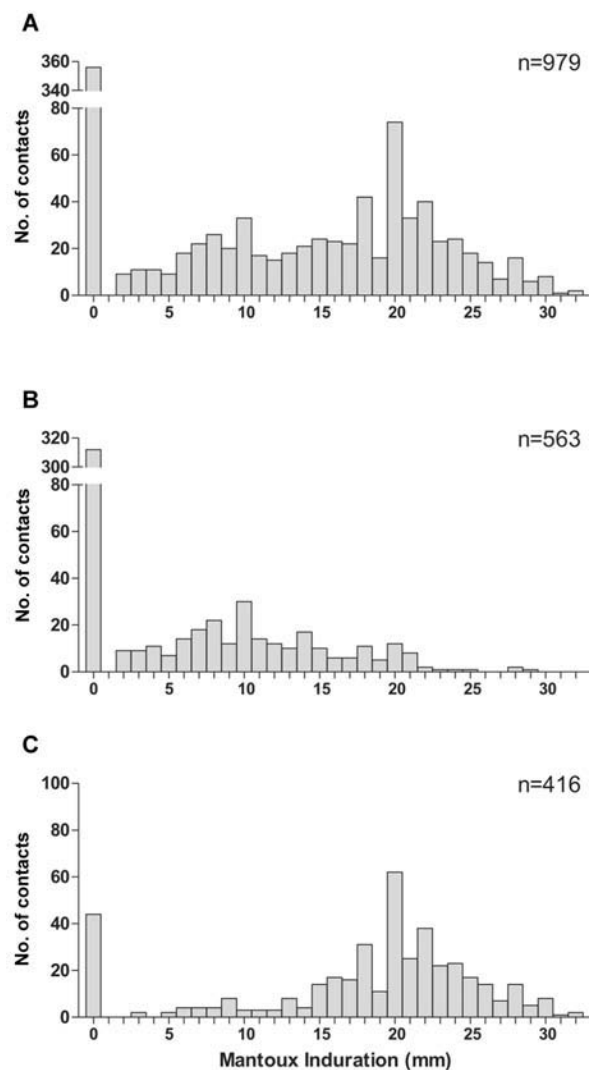


Figure 1. Distribution of tuberculin skin test (TST) indurations, stratified by results of enzyme-linked immunospot assay (ELISPOT). *A*, TST indurations for all child tuberculosis contacts ($n = 979$). *B*, TST indurations for ELISPOT-negative contacts. *C*, TST indurations for ELISPOT-positive contacts. Stratification by ELISPOT results revealed that the bimodal nature of the overall distribution of TST indurations arises through combination of 2 distinct distributions with different modal values, reflecting the infected and uninfected subgroups.

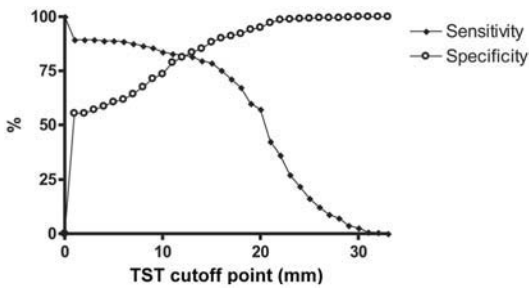


Figure 2. Sensitivity and specificity of the tuberculin skin test (TST) for different cutoff points, with enzyme-linked immunospot assay (ELISPOT) as the reference standard test for latent tuberculosis infection among tuberculosis contacts. *Filled diamonds*, sensitivity; *open circles*, specificity. Sensitivity was calculated as the cumulative proportion of ELISPOT-positive contacts with indurations equal to or larger than the TST cutoff point; specificity was calculated as the cumulative proportion of ELISPOT-negative contacts with indurations smaller than the TST cutoff point.

histories were taken, physical examination and investigations were performed, and demographic data were recorded, as described elsewhere [30]. Complete demographic, clinical, ELISPOT, and TST data were available for 979 of the total of 1024 children enrolled, as described elsewhere [30].

TST. The TST was administered by the Mantoux method, as described elsewhere [30]. For interpretation and analysis of TST induration, 3 different universal cutoff points of induration were used to define a positive result of TST: ≥ 5 mm, ≥ 10 mm, and ≥ 15 mm.

Ex vivo IFN- γ ELISPOT. ELISPOT was performed as described elsewhere [30, 32]. This assay has subsequently been developed into the regulatory agency-approved, commercially available T-SPOT.TB assay (Oxford Immunotec), which uses the same ESAT-6 and CFP-10 peptides. Our predefined cutoff point is the standard used in all previous studies based on this assay, amounting to 9 studies and 1916 participants [16, 17, 19, 20, 22, 29, 32, 38, 39]. The specificity and sensitivity of this assay are described in these studies [16, 17, 19, 20, 22, 29, 32, 38, 39] and in recent reviews [13, 14, 34, 40].

Statistical methods. Determinants of disease prevalence and test sensitivity and specificity were identified using logistic regression modeling, and the significance of each factor was assessed using likelihood-ratio and Wald tests. Observed receiver operating characteristic (ROC) curves were plotted for each subgroup, the areas under the ROC curves were estimated using the trapezoidal rule, and the significance of the differences in areas under the ROC curves was tested.

Smoothed ROC curves were constructed to minimize random error and to remove digit preference, before optimal cutoff points (OCPs) were identified. ROC curves were fitted using the maximum-likelihood latent-scale binormal model of Dorf-

man and Alf [41], and fitted values of sensitivity and specificity were obtained at each observed cutoff point. Estimates of the probability of false-negative diagnosis (Prob(FN)) and the probability of false-positive diagnosis (Prob(FP)) that accounted for disease prevalence were computed for each cutoff point as follows:

$$\text{Prob(FN)} = (1 - \text{sensitivity}) \times \text{prevalence}$$

$$\text{Prob(FP)} = (1 - \text{specificity}) \times (1 - \text{prevalence}) .$$

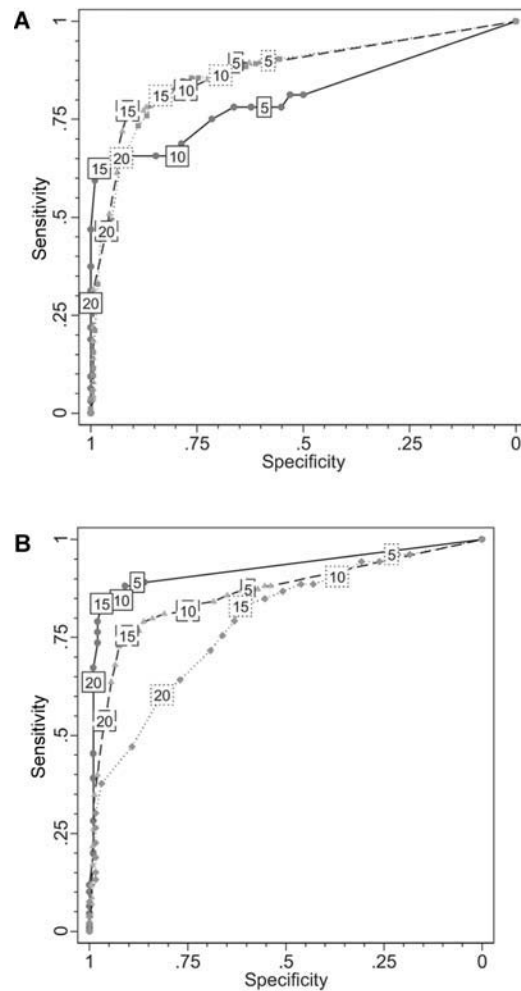


Figure 3. Observed receiver operating characteristic (ROC) curves of tuberculin skin test (TST) performance, compared with that of enzyme-linked immunospot assay. *A*, Observed ROC curves stratified by age group. *Solid line*, children aged <2 years ($n = 130$); *dashed line*, children aged 2–5 years ($n = 275$); *dotted line*, children aged 6–16 years ($n = 574$). *B*, Observed ROC curves stratified by bacille Calmette-Guérin (BCG) vaccination status. *Solid line*, children with no BCG scar ($n = 209$); *dashed line*, children with 1 BCG scar ($n = 652$); *dotted line*, children with >1 BCG scar ($n = 118$). Values in boxes indicate TST thresholds in millimeters.

Table 1. Sensitivity and specificity of the tuberculin skin test (TST), with enzyme-linked immunospot assay as a reference standard, stratified by age and bacille Calmette-Guérin (BCG) vaccination status.

TST cutoff point, stratification and/or group	Sensitivity, % (95% CI)	No. of children	<i>P</i>	Specificity, % (95% CI)	No. of children	<i>P</i>
>5 mm						
All participants	88.9 (85.5–91.8)	416		60.6 (56.4–64.6)	563	
Age group, years						
<2	78.1 (60.0–90.7)	32		59.2 (48.8–69.0)	98	
2–5	89.5 (82.3–94.4)	114	.099 ^a	65.8 (58.0–73.1)	161	.282 ^a
6–16	90.0 (85.8–93.3)	270	.051 ^a	58.2 (52.5–63.8)	304	.867 ^a
2–5 and 6–16 combined	89.8 (86.4–92.7)	384	.049 ^a	60.9 (56.3–65.3)	465	.758 ^a
BCG vaccination status						
No BCG scar	89.1 (81.7–94.2)	110		88.9 (81.0–94.3)	99	
1 BCG scar	87.4 (82.6–91.2)	253	.641 ^b	59.7 (54.7–64.5)	399	<.001 ^b
>1 BCG scar	96.2 (87.0–99.5)	53	.146 ^b	23.1 (13.5–35.2)	65	<.001 ^b
>10 mm						
All participants	83.7 (79.7–87.1)	416		73.5 (69.7–77.1)	563	
Age group, years						
<2	65.6 (46.8–81.4)	32		79.6 (70.3–87.1)	98	
2–5	82.5 (74.2–88.9)	114	.044 ^a	77.6 (70.4–83.8)	161	.711 ^a
6–16	86.3 (81.6–90.2)	270	.004 ^a	69.4 (63.9–74.5)	304	.053 ^a
2–5 and 6–16 combined	85.2 (81.2–88.6)	384	.006 ^a	72.3 (67.9–76.3)	465	.137 ^a
BCG vaccination status						
No BCG scar	84.5 (76.4–90.7)	110		91.9 (84.7–96.4)	99	
1 BCG scar	81.8 (76.5–86.4)	253	.529 ^b	74.9 (70.4–79.1)	399	.001 ^b
>1 BCG scar	90.6 (79.3–96.9)	53	.297 ^b	36.9 (25.3–49.8)	65	<.001 ^b
>15 mm						
All participants	78.6 (74.3–82.5)	416		88.3 (85.3–90.8)	563	
Age group, years						
<2	62.5 (43.7–78.9)	32		98.0 (92.8–99.8)	98	
2–5	77.2 (68.3–84.5)	114	.098 ^a	91.3 (85.8–95.2)	161	.048 ^a
6–16	81.1 (75.9–85.6)	270	.014 ^a	83.6 (78.9–87.5)	304	.002 ^a
2–5 and 6–16 combined	79.9 (75.6–83.8)	384	.024 ^a	86.2 (82.8–89.2)	465	.005 ^a
BCG vaccination status						
No BCG	83.6 (75.4–90.0)	110		97.0 (91.4–99.4)	99	
1 BCG scar	75.5 (69.7–80.7)	253	.088 ^b	90.5 (87.2–93.2)	399	.047 ^b
>1 BCG scar	83.0 (70.2–91.9)	53	.921 ^b	61.5 (48.6–73.3)	65	<.001 ^b

^a *P* value for comparison with group aged <2 years.

^b *P* value for comparison with group with no BCG scar.

OCPs were computed using prevalence of infection estimates obtained from the logistic regression model for each subgroup.

The criteria for the first OCP minimized total diagnostic error, identifying the cutoff point at which the sum of the probabilities of false-positive and false-negative diagnoses of latent tuberculosis infection was lowest, which allowed prevalence of infection to be taken into account. We also report the range of cutoff points for which the probability of diagnostic error was within 2% of this minimum value, to describe the sensitivity of overall performance to cutoff-point selection. A second OCP was defined as the cutoff point at which the consequences of diagnostic error were minimized. Although it is

clear that the consequences of false-negative diagnoses (i.e., missing a case of latent tuberculosis infection) are greater than those of false-positive diagnoses (i.e., incorrectly diagnosing latent tuberculosis infection and giving unnecessary treatment), the relative disutility of these consequences is not known. We created a disutility score as follows:

$$\text{Disutility score} = \text{Prob}(\text{FP}) + [k \times \text{Prob}(\text{FN})],$$

where *k* was given values of 2, 5, and 10, corresponding to assumptions of the consequences of false-negative results being 2, 5, and 10 times more catastrophic than the consequences of

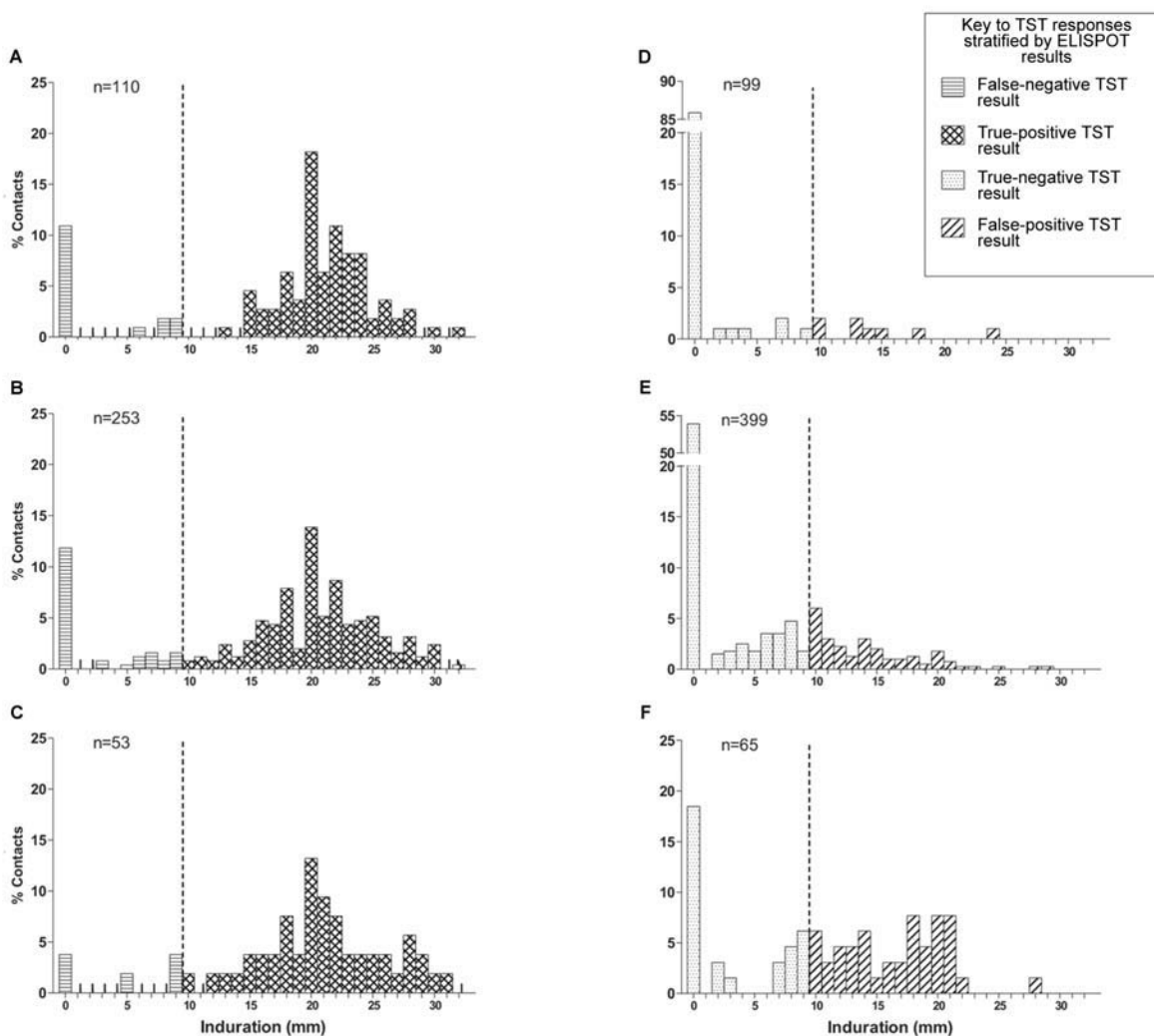


Figure 4. Distribution of tuberculin skin test (TST) indurations, stratified by results of enzyme-linked immunospot assay (ELISPOT) and bacille Calmette-Guérin (BCG) vaccination status. *A*, *B*, and *C*, TST indurations among ELISPOT-positive children. *D*, *E*, and *F*, TST indurations among ELISPOT-negative children. *A* and *D*, TST indurations among children who were unvaccinated. *B* and *E*, TST indurations among children with 1 BCG scar. *C* and *F*, TST indurations among children with >1 BCG scar. *Dashed line*, uniform TST cutoff point of ≥ 10 mm. Distributions of TST indurations among ELISPOT-positive children (*A*, *B*, and *C*) were similar, indicating that an increased number of BCG scars does not affect the sensitivity of the TST. In contrast, results for ELISPOT-negative children (*D*, *E*, and *F*) revealed a substantial decrease in the percentage of children with 0-mm indurations and an increase in the percentage of children with indurations ≥ 10 mm with an increasing number of BCG scars, resulting in lower specificity. The percentage of ELISPOT-negative contacts with TST indurations >0 mm who had no BCG scar, 1 scar, and >1 scar were 14%, 46%, and 82%, respectively ($\chi^2_{\text{trend}} = 73.3$; $P < .001$). The percentage of ELISPOT-negative contacts with positive results of TST as defined by the cutoff point of ≥ 10 mm, were 8%, 25%, and 63% for children with no BCG scar, 1 scar, and >1 scar, respectively ($\chi^2_{\text{trend}} = 59.3$; $P < .001$).

false-positive results, and we identified the cutoff points at which the score was minimized. Results are presented for a k value of 2. We report ranges of cutoff points with disutility values within 2% (relative to the maximum theoretical score) of the OCP score.

Analyses were undertaken using Stata, version 9.0 (Stata Corporation), with the `roccomp` and `rocfite` commands, and using Excel (Microsoft).

RESULTS

Demographic and clinical characteristics of study participants. The median age of the 979 child tuberculosis contacts was 7 years (interquartile range [IQR], 3–11 years), 50.2% were male, and the average number of contacts per household was 2.5. A total of 770 contacts (78.7%) were vaccinated with the BCG vaccine, on the basis of presence or

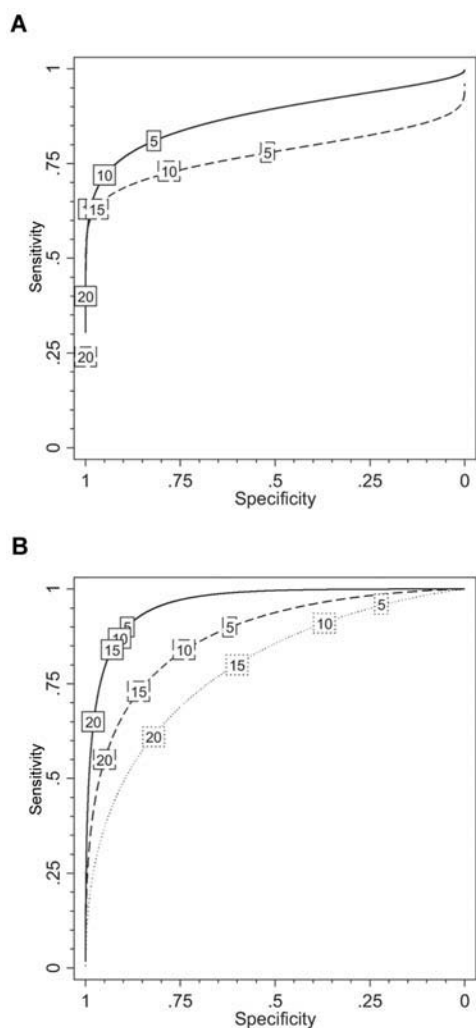


Figure 5. Fitted receiver operating characteristic curves of tuberculin skin test (TST) performance, compared with results of enzyme-linked immunospot assay, stratified by bacille Calmette-Guérin (BCG) vaccination status among patients aged <2 years (A) and 2–16 years (B). Solid lines, children with no BCG scar; dashed lines, children with 1 BCG scar; dotted lines, children with >1 BCG scar. Values in boxes indicate TST thresholds in millimeters.

absence of a BCG scar, including 115 contacts with 2 scars and 3 contacts with 3 scars.

TST and ELISPOT results. On the basis of positive results of ELISPOT, 416 (42.5%) of the children were infected with *M. tuberculosis*. The percentages of TST-positive children were 60.5%, 50.8%, and 40.1% with use of the cutoff points ≥ 5 mm, ≥ 10 mm, and ≥ 15 mm, respectively.

TST indurations for ELISPOT-negative children (figure 1B) were smaller than those for ELISPOT-positive children (figure 1C) (median size [IQR], 0 mm [0–10 mm] vs. 20 mm [16–23 mm]; $P < .001$, by Mann-Whitney U test).

Sensitivity and specificity of the TST, relative to ELISPOT. With a cutoff point of ≥ 1 mm, the relative specificity of the

TST remained >50%, and specificity reached a plateau in the cutoff range 25–30 mm. The relative sensitivity of the TST started at 89.4% and decreased gradually to 78.6% at the cutoff of 15 mm and more steeply with greater cutoffs thereafter (figure 2).

Impact of age on sensitivity and specificity of the TST, relative to ELISPOT. The shape of the ROC curve for children aged <2 years was different from that for older children (figure 3A). With a cutoff point of ≥ 15 mm, the relative sensitivity of the TST decreased from 80% for children aged 2–16 years to 63% for children aged <2 years ($P = .024$) and from 85% to 66%, respectively, with a cutoff point of ≥ 10 mm ($P = .006$) (table 1). The cutoff point of ≥ 5 mm yielded the best relative sensitivity of 78.1% for children aged <2 years, which was significantly lower than that for older children ($P = .049$), at a cost of a very low relative specificity of 59.2% (table 1). Thus, at all cutoff points, the relative sensitivity of the TST was lower for children aged <2 years, and even with the cutoff point of ≥ 5 mm, 7 infected infants would have been missed, because of false-negative results of TST.

Table 1 shows a trend toward lower relative specificity for older children, which remained significant after adjustment for BCG vaccination status ($P = .013$, by test for trend across age categories with a cutoff point of ≥ 15 mm). Despite the strong relationships of relative TST sensitivity and specificity with age, the area under the ROC curve, a measure of overall test accuracy, did not significantly differ among the 3 age groups (0.81, 0.87, and 0.87 for children aged <2 years, 2–5 years, and 6–16 years, respectively; $P = .54$, by test for difference), as can occur when the relationships act in contrasting directions.

Impact of BCG on sensitivity and specificity of the TST, relative to ELISPOT. The area under the ROC curve became significantly smaller with an increasing number of BCG scars (0.93, 0.86, and 0.79 for 0 BCG scars, 1 BCG scar, and >1 BCG scar, respectively; $P = .003$, by test for difference) (figure 3B), indicating a reduction in overall test accuracy. The number of BCG scars significantly adversely affected TST specificity at all cutoff points but did not affect sensitivity (table 1). To investigate this further, we stratified the distribution of TST indurations by ELISPOT results and the number of BCG scars (figure 4).

Optimization of TST cutoff points to minimize diagnostic error, relative to ELISPOT. OCPs that minimized the sum of the probabilities of false-positive and false-negative diagnoses, with prevalence of infection taken into account, were identified from smoothed ROC curves generated for groups defined by age (aged <2 years vs. 2–16 years) and BCG scar status (0 scars vs. 1 scar vs. >1 scar) (figure 5 and table 2). The age groups of 2–5 years and 6–16 years were combined because there was no significant difference in prevalence of infection or TST performance between the groups (table 1 and figure 3A).

Table 2. Optimal cutoff points for the tuberculin skin test (TST) for the study population, stratified by age and bacille Calmette-Guérin (BCG) vaccination status.

Age group (years) and no. of BCG scars	No. of patients	Prevalence, %	Minimize total error rate					Minimize dissimilarity score with the cost of a false-negative result twice the cost of a false-positive result								
			Optimal cutoff, ^a mm	Cutoff range for error $\leq 2\%$, ^b mm	Sensitivity, %	Specificity, %	NPV, %	Error rate of cutoff, %	Optimal cutoff, ^a mm	Cutoff range for error $\leq 2\%$, ^b mm	Sensitivity, %	Specificity, %	NPV, %	Error rate of cutoff, %		
<2																
0	33	33	13	10-15	69	97	91	86 ^c	13	10	7-13	72	95	87	87	13
1	97	22	16	14-17	62	98	91	90 ^c	10	16	13-16	62	98	73	91	10
2-16																
0	176	56	2	2-15	91	88	90	89	10	2	2-10	91	88	90	89	10
1	555	42	15	11-18	73	86	79	81	20	10	2-15	84	74	70	86	22
>1	118	42	20	18-21	61	82	72	74	27	16	11-19	78	63	61	80	31

NOTE. TST performance was most accurate in unvaccinated children aged 2-16 years, for whom ~9 of every 10 diagnoses will be correct. Performance was slightly worse for BCG-naive children aged <2 years, for whom 7 of 8 diagnoses will be correct. In vaccinated children aged 2-16 years, the TST was erroneous in 1 of every 5 diagnoses for patients with 1 BCG scar, and 1 in every 3 diagnoses for patients with >1 BCG scar. NPV, negative predictive value; PPV, positive predictive value.

^a The results are based on the observed prevalence of infection in the 5 population subgroups. Optimal cutoff points were also computed for 3 arbitrary prevalences of infection: 5%, 20%, and 40%. At an infection prevalence of 40%, the optimal cutoff points were lower than at lower levels of prevalence. For children aged <2 years, optimal cutoff points appeared relatively unaffected by infection prevalence; the cutoff point of ≥ 15 mm was robust except for unvaccinated children when disease prevalence was 40%, in which case a lower cutoff point of ≥ 10 mm had the lowest error rate (16%). The selection of cutoff points for children aged ≥ 2 years was more directly affected by prevalence. For unvaccinated children, cutoff points of ≥ 20 , ≥ 15 , and ≥ 10 mm were within a 2% margin of error of the optimal computed cutoff points at prevalences of 5%, 20%, and 40%, respectively. Cutoff points were ~5 mm higher for children with 1 BCG scar and ~10 mm higher for children with >1 scar, compared with those for unvaccinated children. Despite the use of optimal cutoff points tailored to different levels of infection prevalence, the overall error rate of the TST, relative to enzyme-linked immunosorbent assay, increased with increasing prevalence of infection among the target population (data not shown).

^b The range of cutoff points on either side of the optimal cutoff point that increased the error rate by a maximum of 2%.

^c Despite the lower sensitivity of the TST for children aged <2 years, the NPV for this age group is similar to that observed for older children, because the prevalence of infection among children aged <2 years is half that among older children.

Table 2 shows the OCPs for prevalences of infection estimated from the observed data. Results are presented for 2 different analytical strategies (see “Statistical methods” in the Materials and Methods). The estimated performance of the TST, relative to ELISPOT, at these OCPs and the range of cutoff points adjacent to the OCPs with similar overall performance is shown in table 2.

Among unvaccinated children aged ≥ 2 years, for whom relative TST specificity and the prevalence of infection are high (table 2), the probability of false-positive results is low. Optimization, therefore, computed a low cutoff point of ≥ 2 mm, to minimize the summed probabilities of false-negative and false-positive results.

Because of the serious adverse clinical consequences of false-negative results for children aged < 2 years, we calculated OCPs for disutility multipliers of $k = 5$ and $k = 10$ (see “Statistical methods” in the Materials and Methods). For vaccinated children aged < 2 years, at the observed population infection prevalence of 22%, OCPs with disutility multipliers of $k = 5$ and $k = 10$ were ≥ 13 mm and ≥ 11 mm, respectively. For unvaccinated children aged < 2 years, a disutility multiplier of $k = 10$ gave an OCP of ≥ 4 mm, which amounted to a cutoff of any observed response (because 4 mm was the smallest observed response in this group); even this OCP yielded a relative diagnostic sensitivity of only 81%. If the disutility multiplier exceeded 10.7, treatment given to all patients would be a preferable strategy than use of this lowest cutoff point, given the observed study prevalence.

However, even with computed OCPs, the overall error rate of the TST, relative to ELISPOT, increased from 10% for unvaccinated children aged ≥ 2 years to $\geq 20\%$ for children with > 1 BCG scar (table 2).

OCPs were also calculated for the 5 groups with use of arbitrary prevalences of 5%, 20%, and 40%. Cutoff-point selection for children aged < 2 years was unaffected by the prevalence of infection, whereas OCP selection for children aged ≥ 2 years was related to prevalence (see table 2 note).

DISCUSSION

The lack of a reference standard for latent tuberculosis infection greatly complicates the interpretation of TST results, which in turn represents a substantial obstacle to improvement of tuberculosis control. Setting cutoff points, therefore, has relied on empirical comparison of TST indurations in presumptively infected and uninfected populations. Such analyses have hitherto been based on population distributions of TST results for patients with active tuberculosis or hypothetical distributions of TST results computed by subtracting the distribution of TST results among unexposed individuals from that among tuberculosis contacts [42, 43]. We used ELISPOT results as a surrogate reference standard for child tuberculosis contacts, to

assess the effect of BCG vaccination and young age on the specificity and sensitivity of the TST and to optimize TST cutoff points.

The lower relative sensitivity of the TST for children aged < 2 years suggests that the delayed-type hypersensitivity response to *M. tuberculosis* infection in infants is weaker than that in older children [44]. However, ELISPOT can detect very low levels of T cell responses to *M. tuberculosis* infection. This explains its high diagnostic sensitivity, relative to the TST, for infants [17, 33] and young children [29], who have immature cellular immune systems, and for HIV-infected individuals [12, 22, 29].

Our analysis to minimize diagnostic error assumed that the consequences of false-positive and false-negative results were equal, but in practice, the clinical consequences of not treating infected children aged < 2 years are more severe than those of treating uninfected children, because of the high risk of progression to tuberculosis and its attendant high morbidity and mortality [1, 45, 46]. To account for how different clinical outcomes affect clinicians’ interpretation of diagnostic test results, we computed an alternative set of OCPs that minimized a disutility score. This weighted the cutoff-point selection process in favor of minimizing false-negative results and generated cutoff points that were equal to or lower than the cutoff points generated with minimization of the diagnostic error score, because relative TST sensitivity was maximized at the cost of relative specificity. Given the imperative for early treatment of tuberculosis infection in very young children, our results suggest that the TST lacks sufficient sensitivity to reliably rule out a diagnosis of tuberculosis infection among household contacts in this age group. Some national guidelines recommend that child contacts aged < 2 years should receive preventive therapy isoniazid on the basis of positive TST results [5]. Our results lend greater support to a policy of giving universal preventative therapy to all child tuberculosis contacts aged < 2 years, regardless of TST results.

For children aged 2–5 years, the sensitivity of the TST, relative to ELISPOT, was not significantly different from that for children aged > 6 years. Therefore, given that children aged ≥ 2 years are at substantially lower risk of primary progression to active disease than are younger children [2] and given that the relative sensitivity of the TST did not increase further with increasing age, 2 years may be a suitable age threshold above which TST results can be used to guide targeting of isoniazid preventive therapy to child contacts.

In unvaccinated children, TST specificity was high but declined progressively as the number of previous BCG vaccinations increased. In ELISPOT-negative children with > 1 scar, indurations of ≥ 20 mm were not uncommon. The recommended cutoff points in Turkey, although slightly different from our OCPs, nonetheless performed within a 2% margin of error,

compared with the OCPs that were computed to minimize disutility where $k = 2$.

The OCPs represent the best possible performance for the TST in this population, with ELISPOT as the reference standard; however, even these gave error rates of 10%–27%. The problem was most pronounced for children with >1 BCG scar. Several countries (e.g., Turkey and Russia) perform a second BCG vaccination in children and use the TST to diagnose latent tuberculosis infection [10]. Previous studies reached contradictory conclusions about the impact of repeated BCG vaccination on the TST [47–49]. Our results indicate that repeated vaccination has a substantial impact on TST induration, which renders interpretation of the TST unreliable, and even the use of OCPs results in a high proportion of false-positive results among child contacts.

In contrast, TST specificity for children aged <2 years was high, despite the close proximity in time to BCG vaccination (table 1). Thus, notwithstanding the poor sensitivity of the TST for children aged <2 years, its high specificity with use of the cutoff point of ≥ 15 mm makes a positive result of TST a reliable marker of tuberculosis infection in very young children.

Specificity of the TST varies across different populations and regions of the world and depends, in part, on the level of environmental mycobacterial exposure, as well as BCG vaccination status [4, 50]. Thus, although our OCPs have direct relevance to clinical practice in Turkey, they cannot be extrapolated to other populations. Our approach to derivation of OCPs, however, is generalizable. Where deployment of IGRAs is not yet possible, testing of sentinel populations by the TST and IGRA would enable tuberculosis control programs to set more-accurate TST cutoff points tailored to the whole target population.

Our approach could improve epidemiological prevalence surveys by minimizing diagnostic error and avoiding the assumptions about TST induration distributions in presumptively infected populations [51] that are inherent in mathematical methods such as mixture analysis and the mirror method.

In contrast to the consistent evidence from countries with low and medium prevalence of tuberculosis of the correlation of IGRA results with tuberculosis exposure, a recent Gambian study that used a variation of our assay, with different thresholds for scoring results, sensitivity, and specificity [52, 53], found poor correlation; thus, additional work is required for settings with a high burden of tuberculosis. Although the ELISPOT assay used in our study represents an improvement on the TST [12, 16–19, 21, 22, 25, 32, 38] and is already recommended by several national guidelines, it is not a perfect test of latent tuberculosis infection, and our OCPs have been based on imperfect diagnoses in a proportion of our population.

Correlation of baseline TST results with subsequent devel-

opment of tuberculosis over 2 years has recently shown that children deemed to have positive results of TST by the locally defined thresholds were not at increased risk of progressing to tuberculosis, compared with children with negative results of TST [54]. These cutoff points were within 2% of the OCP from our cross-sectional analysis. However, children with positive results of TST defined by a 5-mm induration threshold, a cutoff point also within 2% of the OCP from our cross-sectional analysis, were at a higher risk of progression, compared with children with negative results of TST, although this risk did not reach statistical significance [55]. The recently established prognostic value of positive ELISPOT results, together with the lower specificity of the 5-mm threshold means that significantly more children would need to be given preventive therapy with isoniazid on the basis of positive TST results than on the basis of ELISPOT results to prevent a similar number of incident tuberculosis cases [55].

Our findings could inform tuberculosis control policy. The low sensitivity of the TST for children aged <2 years supports the use of preventive therapy with isoniazid regardless of TST results for these household contacts. For children with 2 BCG scars, the specificity of the TST was very poor, even after adjustment of the cutoff point. Given that a second BCG vaccination confers no additional protection against tuberculosis disease [56] or infection [30] yet renders the TST almost uninterpretable, second vaccinations may be impeding tuberculosis control efforts. Finally, our study provides a mechanism through which the scientific advance of T cell–based testing could be used to improve management of childhood tuberculosis infection in resource-limited settings.

Acknowledgments

We dedicate this article to Prof. Mujdat Basaran, who died 6 years ago. He was the founder of the Department of Paediatrics, Marmara University School of Medicine, and contributed to tuberculosis immunology research with many publications. We thank all the children who took part in the study and their parents. We thank Zeliha Arslan and Filiz Ozturk, for referring contacts from the tuberculosis clinics in Istanbul, and Imogen Staveley, for contributing to the reading and analysis of ELISPOT assays. We acknowledge the crucial support of the Istanbul Association for the Fight Against Tuberculosis and the physicians and nurses of the 7 government-run tuberculosis clinics in the Asian side of Istanbul. We also thank the Department of Tuberculosis of the Ministry of Health of Turkey for their support. We are very grateful to Ben Marais, Luca Richeldi, Frank Cobelens, Hans Rieder, Beate Kampman, Susan Liebeshuetz, Paul Newton, Paul Heath, Paul Fine, and John Innes for critical appraisal of the manuscript and helpful discussions.

Financial support. This work was funded by the Wellcome Trust and the UNICEF/United Nations Development Programme/World Bank/World Health Organization Special Programme for Research and Training in Tropical Diseases. The PhD studentship of D.P.S.D. was supported by the Sir Halley Stewart Trust. J.J.D. was funded by a Department of Health Senior Research Scientist in Evidence Synthesis Award. K.A.M. was supported by a Wellcome Trust Prize studentship. A.L. is a Wellcome Senior Research Fellow in Clinical Science and a member of the Wellcome Trust–funded Centre for Respiratory Infection.

Potential conflicts of interest. A.L. is the lead inventor for several

patents underpinning T cell-based diagnosis. The Lalvani ELISPOT was commercialized by an Oxford University spin-off company (T-SPOT.TB; Oxford Immunotec) in which Oxford University and A.L. have minority shares of equity. A.L. acted as nonexecutive director of Oxford Immunotec in 2003–2007. D.P.S.D. and K.A.M. are named inventors on patents relating to T cell-based diagnosis. All other authors: no conflicts.

References

- Gedde-Dahl T. Tuberculous infection in the light of tuberculin matriculation. *Am J Hyg* **1952**; 56:139–214.
- Marais BJ, Gie RP, Schaaf HS, Beyers N, Donald PR, Starke JR. Childhood pulmonary tuberculosis: old wisdom and new challenges. *Am J Respir Crit Care Med* **2006**; 173:1078–90.
- Hsu KH. Thirty years after isoniazid: its impact on tuberculosis in children and adolescents. *JAMA* **1984**; 251:1283–5.
- Wang L, Turner MO, Elwood RK, Schulzer M, FitzGerald JM. A meta-analysis of the effect of bacille Calmette Guérin vaccination on tuberculin skin test measurements. *Thorax* **2002**; 57:804–9.
- National Collaborating Centre for Chronic Conditions. Tuberculosis: clinical diagnosis and management of tuberculosis, and measures for its prevention and control. London: Royal College of Physicians, **2006**.
- American Thoracic Society. Diagnostic standards and classification of tuberculosis in adults and children. *Am J Respir Crit Care Med* **2000**; 161:1376–95.
- Fundação Nacional de Saúde. Tuberculose—guia de vigilância epidemiológica. Brasília, Brazil: Ministério da Saúde, **2002**:1–99.
- South African Department of Health. The South African tuberculosis control programme—practical guidelines. Cape Town, South Africa: Department of Health, **2000**.
- World Health Organization. Treatment of tuberculosis: guidelines for national programmes. Geneva: World Health Organization, **2003**.
- Türkiye Cumhuriyeti Sağlık Bakanlığı Verem Savaşı Daire Başkanlığı. Türkiye’de tüberkülozun kontrolü için başvuru kitabı. Ankara, Turkey: **2003**. Available at: <http://www.istanbulsglik.gov.tr/w/sb/bh/belge/tuberkitabi.pdf>. Accessed 20 December 2008.
- Chauhan LS, Arora VK. Management of pediatric tuberculosis under the revised national tuberculosis control programme. *Indian J Pediatr* **2004**; 71:341–3.
- Richeldi L. An update on the diagnosis of tuberculosis infection. *Am J Respir Crit Care Med* **2006**; 174:736–42.
- Menzies D, Pai M, Comstock G. Meta-analysis: new tests for the diagnosis of latent tuberculosis infection: areas of uncertainty and recommendations for research. *Ann Intern Med* **2007**; 146:340–54.
- Lalvani A. Diagnosing tuberculosis infection in the 21st century: new tools to tackle an old enemy. *Chest* **2007**; 131:1898–906.
- Pai M, Riley LW, Colford JM Jr. Interferon- γ assays in the immunodiagnosis of tuberculosis: a systematic review. *Lancet Infect Dis* **2004**; 4: 761–76.
- Lalvani A, Pathan AA, Durkan H, et al. Enhanced contact tracing and spatial tracking of *Mycobacterium tuberculosis* infection by enumeration of antigen-specific T cells. *Lancet* **2001**; 357:2017–21.
- Richeldi L, Ewer K, Losi M, et al. T cell-based tracking of multidrug resistant tuberculosis infection after brief exposure. *Am J Respir Crit Care Med* **2004**; 170:288–95.
- Shams H, Weis SE, Klucar P, et al. Enzyme-linked immunospot and tuberculin skin testing to detect latent tuberculosis infection. *Am J Respir Crit Care Med* **2005**; 172:1161–8.
- Lalvani A, Pathan AA, McShane H, et al. Rapid detection of *Mycobacterium tuberculosis* infection by enumeration of antigen-specific T cells. *Am J Respir Crit Care Med* **2001**; 163:824–8.
- Pathan AA, Wilkinson KA, Klenerman P, et al. Direct ex vivo analysis of antigen-specific IFN- γ -secreting CD4 T cells in *Mycobacterium tuberculosis*-infected individuals: associations with clinical disease state and effect of treatment. *J Immunol* **2001**; 167:5217–25.
- Meier T, Eulenbruch HP, Wrighton-Smith P, Enders G, Regnath T. Sensitivity of a new commercial enzyme-linked immunospot assay (T-SPOT-TB) for diagnosis of tuberculosis in clinical practice. *Eur J Clin Microbiol Infect Dis* **2005**; 24:529–36.
- Chapman AL, Munkanta M, Wilkinson KA, et al. Rapid detection of active and latent tuberculosis infection in HIV-positive individuals by enumeration of *Mycobacterium tuberculosis*-specific T cells. *AIDS* **2002**; 16:2285–93.
- Zellweger JP, Zellweger A, Ansermet S, de Senarclens B, Wrighton-Smith P. Contact tracing using a new T-cell-based test: better correlation with tuberculosis exposure than the tuberculin skin test. *Int J Tuberc Lung Dis* **2005**; 9:1242–7.
- Arend SM, Thijsen SF, Leyten EM, et al. Comparison of two interferon- γ assays and tuberculin skin test for tracing tuberculosis contacts. *Am J Respir Crit Care Med* **2007**; 175:618–27.
- Ferrara G, Losi M, D’Amico R, et al. Use in routine clinical practice of two commercial blood tests for diagnosis of infection with *Mycobacterium tuberculosis*: a prospective study. *Lancet* **2006**; 367:1328–34.
- Porsa E, Cheng L, Graviss EA. Comparison of an ESAT-6/CFP-10 peptide-based enzyme-linked immunospot assay to a tuberculin skin test for screening of a population at moderate risk of contracting tuberculosis. *Clin Vaccine Immunol* **2007**; 14:714–9.
- Kang YA, Lee HW, Hwang SS, et al. Usefulness of whole-blood interferon- γ assay and interferon- γ enzyme-linked immunospot assay in the diagnosis of active pulmonary tuberculosis. *Chest* **2007**; 132:959–65.
- Lee JY, Choi HJ, Park IN, et al. Comparison of two commercial interferon- γ assays for diagnosing *Mycobacterium tuberculosis* infection. *Eur Respir J* **2006**; 28:24–30.
- Liebeschuetz S, Bamber S, Ewer K, Deeks J, Pathan AA, Lalvani A. Diagnosis of tuberculosis in South African children with a T-cell-based assay: a prospective cohort study. *Lancet* **2004**; 364:2196–203.
- Soysal A, Millington KA, Bakir M, et al. Effect of BCG vaccination on risk of *Mycobacterium tuberculosis* infection in children with household tuberculosis contact: a prospective community-based study. *Lancet* **2005**; 366:1443–51.
- Detjen AK, Keil T, Roll S, et al. Interferon- γ release assays improve the diagnosis of tuberculosis and nontuberculous mycobacterial disease in children in a country with a low incidence of tuberculosis. *Clin Infect Dis* **2007**; 45:322–8.
- Ewer K, Deeks J, Alvarez L, et al. Comparison of T-cell-based assay with tuberculin skin test for diagnosis of *Mycobacterium tuberculosis* infection in a school tuberculosis outbreak. *Lancet* **2003**; 361:1168–73.
- Richeldi L, Ewer K, Losi M, et al. T-cell-based diagnosis of neonatal multidrug-resistant latent tuberculosis infection. *Pediatrics* **2007**; 119: e1–5.
- Lalvani A, Millington KA. T cell-based diagnosis of childhood tuberculosis infection. *Curr Opin Infect Dis* **2007**; 20:264–71.
- World Health Organization. Global TB database. Available at: http://www.who.int/tb/country/global_tb_database/en/index.html. Accessed 20 December 2008.
- Turkish Ministry of Health. Reported cases and carriers of AIDS distribution by age and gender. **2005**. Available at: <http://www.saglik.gov.tr/extras/istatistikler/2005geribildirim/Tablo%2021.htm>. Accessed 22 December 2008.
- World Health Organization. Immunization profile—the Republic of Turkey. Available at: http://www.who.int/immunization_monitoring/en/globalsummary/countryprofileselect.cfm. Accessed 20 December 2008.
- Lalvani A, Nagvenkar P, Udadia Z, et al. Enumeration of T cells specific for RD1-encoded antigens suggests a high prevalence of latent *Mycobacterium tuberculosis* infection in healthy urban Indians. *J Infect Dis* **2001**; 183:469–77.
- Dosanjh DP, Hinks TS, Innes JA, et al. Improved diagnostic evaluation of suspected tuberculosis. *Ann Intern Med* **2008**; 148:325–36.
- Pai M, Zwerling A, Menzies D. Systematic review: T-cell-based assays for the diagnosis of latent tuberculosis infection: an update. *Ann Intern Med* **2008**; 149:177–84.
- Dorfman DD, Alf E. Maximum likelihood estimation of parameters of

- signal detection theory—a direct solution. *Psychometrika* **1968**;33:117–24.
42. Berkel GM, Cobelens FGJ, de Vries G, Draayer-Jansen IWE, Borgdorff MW. Tuberculin skin test: estimation of positive and negative predictive values from routine data. *Int J Tuberc Lung Dis* **2005**;9:310–6.
 43. Edwards LB, Acquaviva FA, Livesay VT, Cross FW, Palmer CE. An atlas of sensitivity to tuberculin, PPD-B, and histoplasmin in the United States. *Am Rev Respir Dis* **1969**;99(Suppl):1–132.
 44. Lewinsohn DA, Gennaro ML, Scholvinck L, Lewinsohn DM. Tuberculosis immunology in children: diagnostic and therapeutic challenges and opportunities. *Int J Tuberc Lung Dis* **2004**;8:658–74.
 45. Comstock GW, Livesay VT, Woolpert SF. The prognosis of a positive tuberculin reaction in childhood and adolescence. *Am J Epidemiol* **1974**;99:131–8.
 46. Starke JR. Childhood tuberculosis during the 1990s. *Pediatr Rev* **1992**;13:343–53.
 47. Murray DL, Brewer TF. Multiple bacille Calmette-Guérin vaccinations and interpretation of the Mantoux test in children. *Clin Infect Dis* **1995**;21:1533–4.
 48. Shaw LW. Field studies on immunization against tuberculosis. I. Tuberculin allergy following BCG vaccination of school children in Muscogee County, Georgia. *Public Health Rep* **1951**;66:1415–6.
 49. Marcus JH, Khassis Y. The tuberculin sensitivity in BCG vaccinated infants and children in Israel. *Acta Tuberc Pneumol Scand* **1965**;46:113–22.
 50. Black GF, Weir RE, Floyd S, et al. BCG-induced increase in interferon-gamma response to mycobacterial antigens and efficacy of BCG vaccination in Malawi and the UK: two randomised controlled studies. *Lancet* **2002**;359:1393–401.
 51. Dye C, Bassili A, Bierrenbach AL, et al. Measuring tuberculosis burden, trends, and the impact of control programmes. *Lancet Infect Dis* **2008**;8:233–43.
 52. Hill PC, Brookes RH, Fox A, et al. Large-scale evaluation of enzyme-linked immunospot assay and skin test for diagnosis of *Mycobacterium tuberculosis* infection against a gradient of exposure in The Gambia. *Clin Infect Dis* **2004**;38:966–73.
 53. Hill PC, Brookes RH, Adetifa IM, et al. Comparison of enzyme-linked immunospot assay and tuberculin skin test in healthy children exposed to *Mycobacterium tuberculosis*. *Pediatrics* **2006**;117:1542–8.
 54. Lalvani A. Prognostic value of a T-cell-based, interferon- γ biomarker in Turkish children—authors' response. *Ann Intern Med* **2008**; published online 7 November. Available at: <http://www.annals.org/cgi/eletters/149/11/777>. Accessed 22 December 2008.
 55. Bakir M, Millington KA, Soysal A, et al. Prognostic value of a T-cell-based, interferon- γ biomarker in children with tuberculosis contact. *Ann Intern Med* **2008**;149:777.
 56. Rodrigues LC, Pereira SM, Cunha SS, et al. Effect of BCG revaccination on incidence of tuberculosis in school-aged children in Brazil: the BCG-REVAC cluster-randomised trial. *Lancet* **2005**;366:1290–5.