ORIGINAL ARTICLE

Diagnostic accuracy of interleukin 27 for tuberculous pleural effusion: two prospective studies and one meta-analysis

Wen Wang,[1] Qiong Zhou,[2] Kan Zhai,[1] Yao Wang,[1] Jing-Yuan Liu,[2] Xiao-Juan Wang,[1] Zhen Wang,[1] Jian-Chu Zhang,[2] Zhao-Hui Tong,[1] Huan-Zhong Shi[1]

[1]Department of Respiratory and Critical Care Medicine, Beijing Institute of Respiratory Medicine and Beijing Chaoyang Hospital, Capital Medical University, Beijing, China
[2]Department of Respiratory and Critical Care Medicine, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Correspondence to
Dr Huan-Zhong Shi, Department of Respiratory and Critical Care Medicine, Beijing Chaoyang Hospital, Capital Medical University, 8 Gongti Nanlu, Chaoyang District, Beijing 100020, China; shihuanzhong@sina.com

WW and QZ contributed equally.

► http://dx.doi.org/10.1136/thoraxjnl-2017-210918

ABSTRACT
Background Accurate differentiating diagnosis is essential for choosing treatment for exudative pleural effusions.
Objective To establish the diagnostic accuracy of interleukin 27 for tuberculous pleural effusion (TPE).
Methods First, the concentrations of pleural interleukin 27, interferon-gamma and adenosine deaminase were compared between 51 patients with TPE and 103 with non-TPEs (Beijing cohort), and their diagnostic values were evaluated. These were further verified in another independent population (Wuhan cohort, n=120). In the second part of the study, we performed a meta-analysis.
Results With a cut-off value of 591.4 ng/L in the Beijing cohort, the area under the curve, sensitivity, specificity, positive predictive value and negative predictive value of interleukin 27 to diagnose TPE were 0.983 (95% CI 0.947 to 0.997), 96.1% (86.5% to 99.5%), 99.0% (94.7% to 100%), 98.0 (89.4 to 99.9) and 98.1 (93.3 to 99.8), respectively. Excellent diagnostic accuracy of interleukin 27 was also found in the Wuhan cohort and was further confirmed in the meta-analysis. The diagnostic performance of interleukin 27 was comparable to that of interferon-gamma and was more accurate than that of adenosine deaminase. Since the post-test probability of a negative result was always <0.1%, a negative test was considered to exclude TPE in all tuberculosis prevalence settings.
Conclusions Interleukin 27 can be used to diagnose TPE in a high prevalence setting, and a negative result can also be reliably used to rule out TPE in all prevalence settings.

Key messages

What is the key question?
► Can tuberculous pleural effusion be distinguished from those pleural effusions with the other aetiologies using interleukin 27 concentrations in pleural fluid?

What is the bottom line?
► In this study we show that interleukin 27 is a valuable indicator for ruling in tuberculous pleural effusion in a high prevalence setting, and for ruling out tuberculous pleural effusion in all prevalence settings.

Why read on?
► A design including two prospective independent cohorts and a meta-analysis allowed this study to establish the exact diagnostic accuracy of interleukin 27 for tuberculous pleural effusion.

INTRODUCTION
Although it is curable, tuberculosis remains one of the most common causes of pleural effusions globally, especially in the developing countries.[1] Tuberculous pleural effusion (TPE) is one of the most common forms of extrapulmonary tuberculosis.[2] The definitive diagnosis of TPE depends on the demonstration of *Mycobacterium tuberculosis* in the sputum, pleural fluid or pleural biopsy specimens.[3] A presumptive diagnosis can also be established with reasonable certainty by demonstrating granuloma in the parietal pleura[1–3] or elevated concentrations of adenosine deaminase (ADA)[4] or interferon-γ (IFN-γ)[5] in pleural fluid.

A few years ago, we reported for the first time that interleukin 27 (IL-27)—a cytokine of the IL-12 cytokine family produced by antigen-presenting cells under the stimulation of pathogen-associated molecular patterns binding to their toll-like receptors—can distinguish between TPE and non-TPEs, with a sensitivity of 92.7% and a specificity of 99.1%.[6] In another study we indicated that the diagnostic efficiency of IL-27 is better than that of ADA.[7] On the other hand, one study from a country with a low tuberculosis burden showed that IL-27 is less efficient than ADA,[8] and two studies showed that the product of IL-27 and ADA improves the diagnostic sensitivity of ADA.[8 9] We therefore performed these prospective studies, with two independent replication cohorts and a meta-analysis, to establish the precise role of IL-27 for the diagnosis of TPE.

METHODS
Study populations
The study was approved by the ethics committees of both participating institutions, and all study participants provided written informed consent. Between January 2014 and June 2015, all consecutive adult patients with pleural effusions admitted to the Department of Respiratory and Critical Care Medicine, Beijing Chaoyang Hospital, Capital Medical University, Beijing (Beijing cohort) were

Table 1 Demographic, cytological and biochemical characteristics in pleural effusions*

Variable	Beijing cohort					Wuhan cohort					p Value† 0.826
	TPE (n=51)	MPE (n=61)	PPE (n=28)	Miscellaneous (n=14)	p Value‡	TPE (n=44)	MPE (n=47)	PPE (n=19)	Miscellaneous (n=10)	p Value‡	
Sex, male, %	64.7	53.8	64.3	64.3	0.850	65.9	55.3	63.2	70.0	0.694	0.548
Age, years	43.2±2.6	63.2±1.6	61.5±2.4	61.8±4.7	<0.001	45.5±2.7	59.5±1.5	62.2±3.6	57.6±4.4	0.003	0.116
RBC×10⁹/L	2.4 1.6–5.0	3.9 0.9–8.9	2.6 0.8–5.4	1.3 0.5–6.5	0.341	2.3 0.9–6.3	3.6 1.1–8.8	2.5 0.9–6.4	1.5 0.4–5.4	0.389	0.788
Nucleated cell counts×10⁹/L	2.5 1.0–3.9	2.1 1.4–3.5	3.8 2.3–12.1	1.8 0.5–5.4	0.418	2.4 1.3–4.9	1.9 1.2–5.0	4.1 1.8–14.2	1.7 0.8–5.9	0.392	0.694
Protein, g/L	47.3±3.6	42.5±2.7	47.7±2.4	32.6±8.2	0.292	44.4±3.0	46.7±3.1	48.2±2.5	30.6±7.9	0.264	0.749
Glucose, mmol/L	5.8±1.1	5.6±1.2	4.1±0.9	6.1±1.2	0.340	5.7±1.1	5.4±1.0	3.8±0.9	6.0±1.1	0.363	0.868
LDH, U/L	829 421–1826	520 429–1453	1056 613–3620	358 124–745	0.085	750 368–1698	464 254–1266	828 275–1432	304 143–681	0.078	0.493

*Data are presented as means±SEMs or medians (25th – 75th centiles). †Comparisons were performed between the Beijing and Wuhan cohorts using X^2 test, Student's t test, or Mann–Whitney U test. ‡Comparisons were performed using X^2 test, analysis of variance (ANOVA) followed by Bonferroni's test or Kruskal-Wallis ANOVA on ranks.
LDH, lactate dehydrogenase; MPE, malignant pleural effusion; PPE, parapneumonic pleural effusion; RBC, red blood cell; TPE, tuberculous pleural effusion.

screened for this prospective study (see online supplementary figure 1). One hundred and fifty-four patients with an established definite cause of pleural effusion were included in the study. This group was used to evaluate the diagnostic accuracy of IL-27, ADA and IFN-γ for TPE. This was subsequently validated in an independent cohort of 120 consecutive patients with pleural effusions who were recruited prospectively during the same period from the Department of Respiratory and Critical Care Medicine, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan (Wuhan cohort, online supplementary figure 2). The characteristics of both study cohorts are summarised in table 1, and the aetiological distribution of the pleural effusions is presented in online supplementary table 1).

Diagnostic criteria

TPE was diagnosed if Ziehl-Neelsen stains or Lowenstein-Jensen cultures of pleural fluid, sputum or pleural biopsy specimens were positive or if granulomas were present in the parietal pleural biopsy specimens (online supplementary table 2). A diagnosis of malignant pleural effusion was established by demonstration of malignant cells in pleural fluid and/or on pleural biopsy specimens. A parapneumonic pleural effusion was diagnosed as any effusion associated with bacterial pneumonia, lung abscess or bronchiectasis; empyema was also included in this group. The remaining effusions comprised miscellaneous pleural effusions.

The patients recruited to this study were those with definite known causes contributing to their pleural effusions. At the time of sample collection, none of the patients had received any anticancer treatment, antituberculosis therapy, corticosteroids or other non-steroidal anti-inflammatory drugs.

Measurement of IL-27, ADA and IFN-γ

Pleural fluid was collected by diagnostic thoracentesis before the patient had received any treatment and was rapidly transferred to the laboratory with a blood sample obtained simultaneously from the same patient. Pleural fluid and blood samples were centrifuged at 1500 rpm for 10 min at 4°C, and the supernatants were aliquoted and stored at −80°C awaiting measurement of IL-27, ADA and IFN-γ.

The concentrations of IL-27 and IFN-γ in pleural fluid and serum were measured by ELISA kits (eBioscience, Inc, San Diego, California, USA) according to the manufacturer's protocols. ADA activity was determined using colorimetric method kits (InTec Products, Inc, Xiamen, China) in accordance with the manufacturer's instructions. The technologists running the assays were blinded to the nature of the samples, and the code was broken to the statisticians after construction of the database.

Statistical analysis

The sample sizes were calculated based on the assumption of a sensitivity of 92.7% and a prevalence of TPE of about 40% obtained in our previous study.[6] With a two-sided α of 5%, we aimed to recruit about 130 patients in both the Beijing and Wuhan cohort to obtain a precision of 7% in the estimation of the sensitivity of pleural IL-27 in the diagnosis of TPE.

The concentrations of IL-27 and the activities of ADA were normally distributed, as determined by the Kolmogorov-Smirnov test, and are presented as means±SEMs. IFN-γ data are presented as medians (25th to 75th centiles) since they were not normally distributed. Differences between two or multiple groups were compared using Student's t test, Mann–Whitney U test, one-way analysis of various (ANOVA) or Kruskal-Wallis ANOVA on ranks, as appropriate. All tests were two tailed and the p value for multiple testing was adjusted by Bonferroni's test. Comparisons of data in the pleural effusion and in the corresponding serum were made using a paired t test or Wilcoxon signed-rank test, as appropriate.

Receiver operating characteristic curves were drawn, and areas under the curves (AUCs) were calculated to determine the diagnostic value of the concentrations of each biomarker in the pleural effusion, including sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, positive predictive value and negative predictive value[10 11]; AUCs were compared using the z statistic with the Hanley and McNeil procedure.[12] The optimum cut-off values were defined based on their maximum Youden index (sensitivity+specificity−1). The parameters of diagnostic accuracy are shown together with their 95% CIs. To verify the diagnostic accuracy of IL-27, ADA and IFN-γ in the Wuhan cohort, cases with the concentration of each biomarker complying with the corresponding cut-off value obtained in the Beijing cohort, were considered as positive results. All statistical analyses were performed with SPSS and MedCalc software, and p<0.05 was considered to indicate statistical significance.

Meta-analysis

Medline using the PubMed interface, Embase and the Web of Science were searched for suitable studies up to 8 November 2016; no early date limit was applied. Search keywords included: 'IL-27/interleukin 27', 'pleurisy/pleural effusion' and 'tuberculous/tuberculosis'. Articles were also identified by use of the related-articles function in PubMed. References of articles identified were searched manually. Two reviewers independently judged study eligibility while screening the publications. A study was included in the meta-analysis when it reported the concentrations of IL-27 in pleural fluid. Data retrieved from the studies included participant characteristics, test methods, sensitivity and specificity data, cut-off values, publication year and methodological quality.

Using two statistical software programs (Stata, version 9; Stata Corporation; College Station, Texas, USA; and Meta-DiSc for Windows; XI Cochrane Colloquium; Barcelona, Spain), the pooled sensitivity, specificity, positive likelihood ratio, negative likelihood ratio and diagnostic OR of IL-27 were computed by standard methods.[13] The analysis was based on a summary receiver operating characteristic curve.[13] [14] The sensitivity and specificity for the single test threshold identified for each study were used to plot a summary receiver operating characteristic curve.[14] A random-effects model was used to calculate the average sensitivity, specificity, positive likelihood ratio, negative likelihood ratio and diagnostic OR across the studies.[15] [16] The X^2 and Fisher exact tests were used to detect statistically significant heterogeneity across studies. Since publication bias is of concern for meta-analysis of diagnostic studies, we tested for the potential presence of this bias using funnel plots and the Egger test.[17]

Post-test probability of TPE was estimated by the formula: post-test probability=post-test odds/(post-test odds+1), where post-test odds = (prevalence/[1 – prevalence])×likelihood ratio.

RESULTS

In this study, pleural effusions without a definite known cause were excluded, as they might have been malignant but missed by cytology. A total of 154 patients in the Beijing cohort (online supplementary figure 1) and 120 in the Wuhan cohort (online supplementary figure 2), who had paired pleural fluid and serum samples for determining IL-27, ADA and IFN-γ, were included. No differences in the demographic, cytological and biochemical characteristics of the pleural effusions were found between the Beijing and Wuhan cohorts (table 1).

Concentrations of IL-27, ADA and IFN-γ in pleural effusions

As expected and consistent with our previous findings,[6] [7] we once again observed that the concentrations of IL-27, ADA and IFN-γ in TPE were significantly higher than in patients in the non-TPE group (Bonferroni's test, all p<0.001), and that there were no differences among malignant, parapneumonic and miscellaneous effusions (Bonferroni's test, all p>0.05) (table 2 and figure 1A, C, E). We also noted that in the patients with TPE concentrations of IL-27, ADA and IFN-γ were all higher in pleural fluid than serum (all p<0.001). In addition, the serum concentrations of IL-27 in patients with TPE did not differ from those in the three non-TPE groups (Bonferroni's test, all p>0.05).

Diagnostic values of IL-27, ADA and IFN-γ in the Beijing cohort

With a cut-off value of 591.4 ng/L in the Beijing cohort, the AUC of IL-27 to differentiate TPE from non-TPEs was 0.983 (95% CI, 0.947 to 0.997; p<0.001) (figure 1B and table 3). Also as shown in table 3, the sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, positive predictive value and negative predictive value of IL-27 were 96.1%, 99.0%, 99.0, 0, 98.0 and 98.1, respectively.

The parameters of diagnostic accuracy of ADA and IFN-γ are also presented in figure 1 and table 3. ADA and IFN-γ had AUCs of 0.900 (0.842 to 0.943) (figure 1D) and 0.973 (0.933 to 0.992) (figure 1F), respectively. A comparison of the receiver operating characteristic curves showed that the AUC of ADA was much lower than that of IL-27, with a difference between the areas of 0.082 (95% CI 0.018 to 0.147; z=2.514, p=0.012). The AUC of ADA was also lower than that of IFN-γ (0.072 (0.010 to 0.135); z=2.287, p=0.022). There was no significant difference in the AUCs between IL-27 and IFN-γ (0.010 (–0.020 to 0.040); z=0.643, p=0.521).

Table 2 Concentrations of IL-27, ADA and IFN-γ in pleural fluid and/or serum according to study cohort*

Variable	Beijing cohort				Wuhan cohort			
	TPE (n=51)	MPE (n=61)	PPE (n=28)	Miscellaneous (n=14)	TPE (n=44)	MPE (n=47)	PPE (n=19)	Miscellaneous (n=10)
IL-27								
PF, ng/L	892.7±35.2†‡	298.3±13.1	270.4±16.1	252.3±12.1	885.3±39.2†	288.7±13.0	306.8±22.6	252.3±18.3
Serum, ng/L	334.5±25.7	326.6±9.2	321.2±21.2	309.9±11.9	314.7±10.0	320.7±10.2	312.6±18.7	279.4±19.1
ADA								
PF, U/L	49.2±3.7†‡	15.7±1.3	14.5±1.6	15.9±2.0	48.2±4.9†	13.3±1.2	14.7±1.1	18.6±2.0
Serum, U/L	13.8±0.8		13.5±0.9	14.2±1.5	15.0±0.9		12.5±0.9	13.9±1.2
IFN-γ								
PF, ng/L	1,169.8§¶ (341.9–3881.6)	29.2 (20.2–39.9)	32.6 (18.5–47.3)	38.2 (26.5–57.4)	1,126.3§ (322.6–3177.0)	24.9 (18.3–38.1)	44.2 (19.9–106.0)	47.7 (32.4–106.6)
Serum, ng/L	27.1§ (14.7–51.3)	12.4 (8.7–15.6)	16.8 (9.7–26.0)	11.8 (8.8–27.2)	30.7§ (12.7–50.5)	11.8 (8.3–13.2)	10.7 (8.3–15.6)	11.2 (10.2–13.8)

*Data are presented as mean ±SEM or median (25th – 75th centile). †p<0.001 compared with each non-TPE group using analysis of variance (ANOVA) followed by Bonferroni's test; ‡p<0.001 compared with the corresponding serum using paired t test; §p<0.001 with each non-TPE group using Kruskal-Wallis ANOVA; ¶p<0.001 compared with the corresponding serum using Wilcoxon signed-rank test; **p<0.05 compared with TPE group using ANOVA followed by Bonferroni's test.
ADA, adenosine deaminase; IFN, interferon; IL, interleukin; MPE, malignant pleural effusion; PF, pleural fluid; PPE, parapneumonic pleural effusion; TPE, tuberculous pleural effusion.
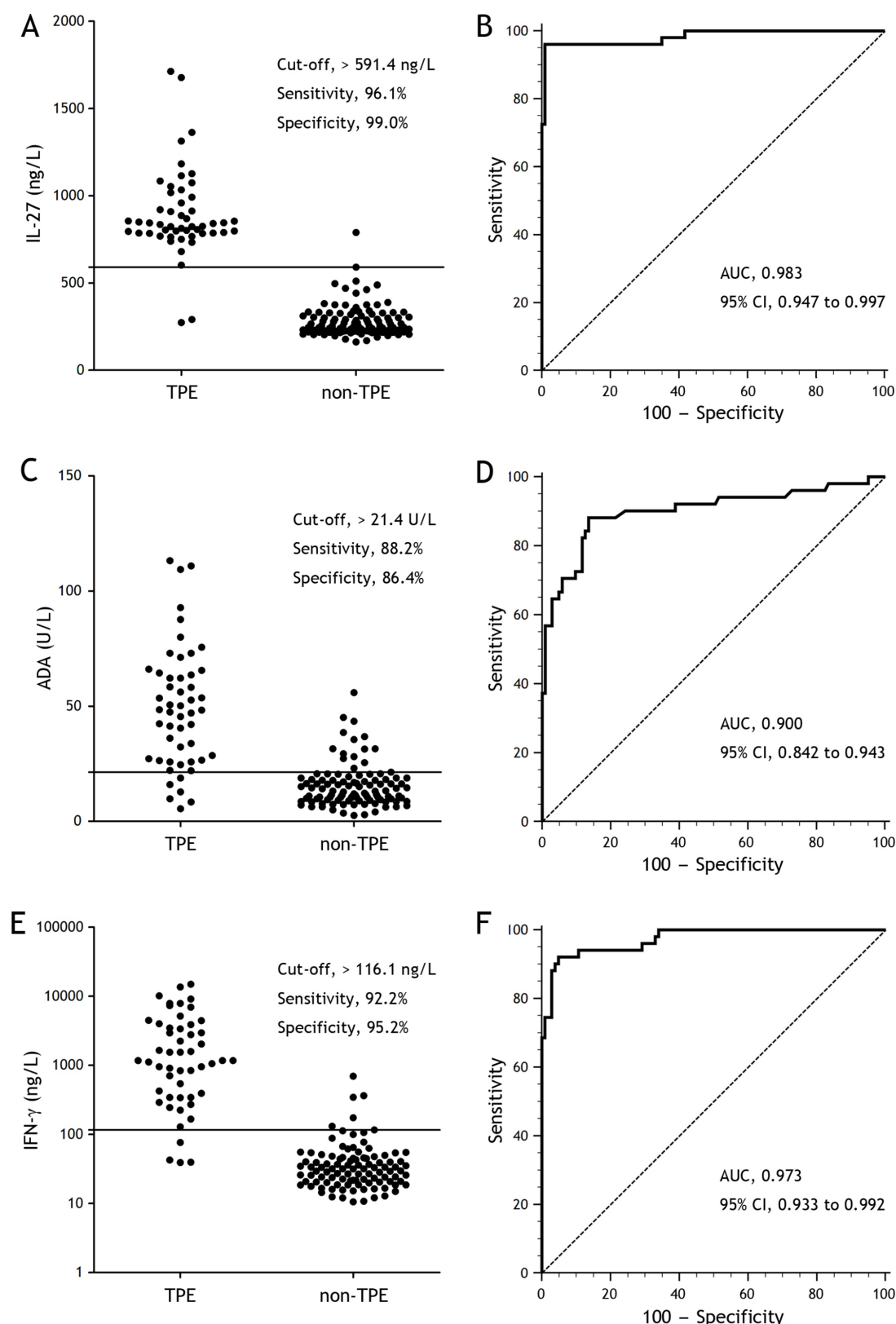
Figure 1 Diagnostic accuracy of interleukin 27 (IL-27), adenosine deaminase (ADA) and interferon-γ (IFN-γ) in distinguishing patients with tuberculous pleural effusion (TPE) from those with non-tuberculous pleural effusions in the Beijing Cohort. (A) A cut-off value of 591.4 ng/L for IL-27 showed high sensitivity and specificity for the detection of TPE. (B) A receiver operating characteristic curve shows the diagnostic performance of the IL-27 assay performed on pleural fluid samples obtained from the Beijing study that included 51 patients with TPE and 103 patients with non-TPE. (C) A cut-off value of 21.4 U/L for ADA showed relatively high sensitivity and specificity for the detection of TPE. (D) A receiver operating characteristic curve shows the diagnostic performance of the ADA assay performed on pleural fluid samples obtained from the Beijing study. (E) A cut-off value of 116.1 ng/L for IFN-γ showed high sensitivity and specificity for the detection of TPE. (F) A receiver operating characteristic curve shows the diagnostic performance of the IFN-γ assay performed on pleural fluid samples obtained from the Beijing study. AUC, area under the curve.

Table 3 Diagnostic performance of pleural IL-27, ADA and IFN-γ in differentiating between patients with tuberculous pleural effusion (TPE) and those with non-TPE in the Beijing cohort (n=154)

Variable	Cut-off value	AUC (95% CI)	Sensitivity (%)	Specificity (%)	PLR	NLR	PPV	NPV
IL-27	591.4 ng/L	0.983* (0.947 to 0.997)	96.1 (86.5 to 99.5)	99.0 (94.7 to 100.0)	99.0 (14.1 to 696.4)	0 (0 to 0.2)	98.0 (89.4 to 99.9)	98.1 (93.3 to 99.8)
ADA	21.4 U/L	0.900 (0.842 to 0.943)	88.2 (76.1 to 95.6)	86.4 (78.2 to 92.4)	6.5 (3.9 to 10.7)	0.1 (0.1 to 0.3)	76.3 (63.4 to 86.4)	93.7 (86.7 to 97.7)
IFN-γ	116.1 ng/L	0.973† (0.933 to 0.992)	92.2 (81.1 to 97.8)	95.2 (89.0 to 98.4)	19.0 (8.0 to 44.8)	0.1 (0 to 0.2)	90.4 (79.0 to 96.8)	96.1 (90.3 to 98.9)
IL-27•ADA	14 496.8 ng U/L^2	0.979‡ (0.943 to 0.995)	90.2 (78.6 to 96.7)	100.0 (96.5 to 100.0)	–	0.1 (0 to 0.2)	100.0 (92.3 to 100.0)	95.4 (89.5 to 98.5)

*p = 0.012, †p=0.022, ‡p<0.001, compared with ADA using the z statistic, respectively.

ADA, adenosine deaminase; AUC, area under the curve; IL-27, interleukin 27; IFN-γ, interferon-γ; IL-27•ADA, product of IL-27 and ADA; NLR, negative likelihood ratio; NPV, negative predictive value; PLR, positive likelihood rati0; PPV, positive predictive value.

It has been reported that the product of IL-27 and ADA is more accurate than ADA for diagnosing TPE,[8 9] but this was not shown in our study. Although a perfect specificity of the product of IL-27 and ADA could be seen (100.0%), it was at a cost of the reduced sensitivity (90.2%). Z-statistical analysis further confirmed that the AUC of the product of IL-27 and ADA was not higher than that of IL-27 (0.003 (–0.025 to 0.031); z=0.227, p=0.820).

Diagnostic values of IL-27, ADA and IFN-γ in the Wuhan cohort

We performed separately another prospective blinded validation study (Wuhan cohort) using the cut-off values of IL-27, ADA and IFN-γ obtained from Beijing cohort. Excellent discrimination was found with IL-27 between 44 patients with TPE and 76 controls in the Wuhan cohort (table 4 and online supplementary figure 3). Similar to the findings from the Beijing cohort, z-statistical analysis showed that the AUC of IL-27 was not higher than that of IFN-γ (0.029 (–0.027 to 0.085); z=1.024, p=0.306), but was significantly higher than that of ADA (0.105 (0.030 to 0.180]; z=2.728, p=0.006). In addition, the AUC of IFN-γ was also higher than that of ADA (0.075 (0.006 to 0.144); z=2.141, p=0.032).

Overall diagnostic accuracy of IL-27 from meta-analysis

After independent review, we found that pleural concentrations of IL-27 had been reported in five publications.[6–9 18] Our study was conducted separately in two different populations, and thus we treated these research data as two independent studies. Thus, seven studies were considered eligible for inclusion in the meta-analysis (online supplementary figure 4), and the characteristics of these studies are shown in online supplementary table 3. The average samples size of the seven pleural IL-27 studies included was 165 (range 76–431); the subjects included 323 patients with TPE and 834 with non-TPEs. In two publications[8 18] and our current two studies, the diagnosis of TPE was made by bacteriological or histological studies or both (gold standard); in the remaining three publications,[6 7 9] the majority of patients with TPE were diagnosed based on bacteriological or histological results, and a minority were diagnosed based on clinical course, including clinical presentation, pleural fluid analysis, radiology and the responsiveness to antituberculosis chemotherapy.

We assessed methodological quality using guidelines published by the QUADAS (quality assessment for studies of diagnostic accuracy, maximum score 14) tool[19] (ie, appraisal by use of empirical evidence, expert opinion and formal consensus to assess the quality of primary studies of diagnostic accuracy) and the STARD (standards for reporting diagnostic accuracy, maximum score 25) initiative[20] (ie, guidelines that aim to improve the quality of reporting in diagnostic studies). As shown in online supplementary table 3, the quality of study design and reporting diagnostic accuracy of most studies were good, as all studies had high QUADAS scores (≥10) and only one study[18] had a low STARD score (<15). In addition, the funnel plots for publication bias showed some asymmetry (online supplementary figure 5), and the Egger test was significant for IL-27 (p=0.011), indicating that there was a potential for publication bias.

Figure 2A shows the forest plot of sensitivity and specificity for seven pleural IL-27 assays in the diagnosis of TPE. The pooled sensitivity, specificity, positive likelihood ratio, negative likelihood ratio and diagnostic OR was 93.8% (95% CI 90.6% to 96.2%), 91.7% (89.6% to 93.5%), 29.87 (7.47 to 119.35), 0.08 (0.05 to 0.12) and 427.86 (101.26 to 1807.80), respectively. Figure 2B shows that the summary receiver operating characteristic curve of IL-27 was positioned very near the desirable upper left corner, and that the maximum joint sensitivity and specificity was 0.930 (SEM 0.026); while the AUC was 0.976 (SEM 0.015).

When the pre-test prevalence of TPE is 10% (a hypothetical patient at low-risk) and when the IL-27 result was positive, the estimated post-test probability was 76.7% (table 5); whereas a negative test result almost certainly excluded TPE (post-test probability, 0.009%). For a patient with a 25% or 50% pre-test risk, a positive test result increased the probability to 90.8% or 96.8%, respectively; while the absence of a high IL-27 concentration reduced the disease probability to 0.025% or 0.074%, respectively.

DISCUSSION

A definite diagnosis of TPE is confirmed only when *Mycobacterium tuberculosis* is demonstrated in sputum or pleural specimens, or when epithelioid cell granulomas and/or caseating granulomas are found in pleural biopsy specimens.[1 2] However, sensitivities of microscopic examination (0–1%)[21–23] and culture (24–58%)[22–25] of pleural fluid for diagnosing TPE are not high owing to the low load of *Mycobacterium tuberculosis*. Therefore, differential diagnosis of TPE sometimes requires more invasive methods, such as medical thoracoscopy, when one or more thoracenteses fail to establish a definite diagnosis. Almost all cases of TPE can be diagnosed by medical thoracoscopy[26]; however, not everyone is suitable for, or willing to undergo, medical thoracoscopy.

Table 4 Diagnostic performance of pleural IL-27, ADA and IFN-γ in in differentiating between patients with tuberculous pleural effusion (TPE) and those with non-TPE in the Wuhan cohort (n=120)

Variable	AUC (95% CI)	Sensitivity (%)	Specificity (%)	PLR	NLR	PPV	NPV
IL-27, >591.4 ng/L	0.982* (0.939 to 0.998)	97.7 (88.0 to 99.9)	98.7 (92.9 to 100.0)	74.3 (10.6 to 520.8)	0 (0 to 0.2)	97.7 (88.0 to 99.9)	98.7 (92.9 to 100.0)
ADA, >21.4 U/L	0.877 (0.805 to 0.930)	88.6 (75.4 to 96.2)	86.8 (77.1 to 93.5)	6.7 (3.7 to 12.1)	0.1 (0.1 to 0.3)	79.6 (65.5 to 89.8)	93.0 (84.3 to 97.7)
IFN-γ, >116.1 ng/L	0.953‡ (0.898 to 0.983)	93.2 (81.3 to 98.6)	97.4 (90.8 to 99.7)	35.4 (9.0 to 139.3)	0.1 (0 to 0.2)	95.3 (84.0 to 99.5)	96.1 (89.0 to 99.2)
IL-27•ADA, >144960.8 U/L²	0.943‡ (0.885 to 0.977)	88.6 (75.4 to 96.2)	100.0 (95.3 to 100.0)	–	0.1 (0.1 to 0.3)	100.0 (90.7 to 100.0)	93.8 (86.2 to 98.0)

*p = 0.006, †p=0.032, 0 p=0.010, compared with ADA using the z statistic, respectively.
ADA, adenosine deaminase; AUC, area under the curve; IL-27, interleukin 27; IFN-γ, interferon-γ; IL-27•ADA, the product of IL-27 and ADA; NLR, negative likelihood ratio; NPV, negative predictive value; PLR, positive likelihood ratio; PPV, positive predictive value.
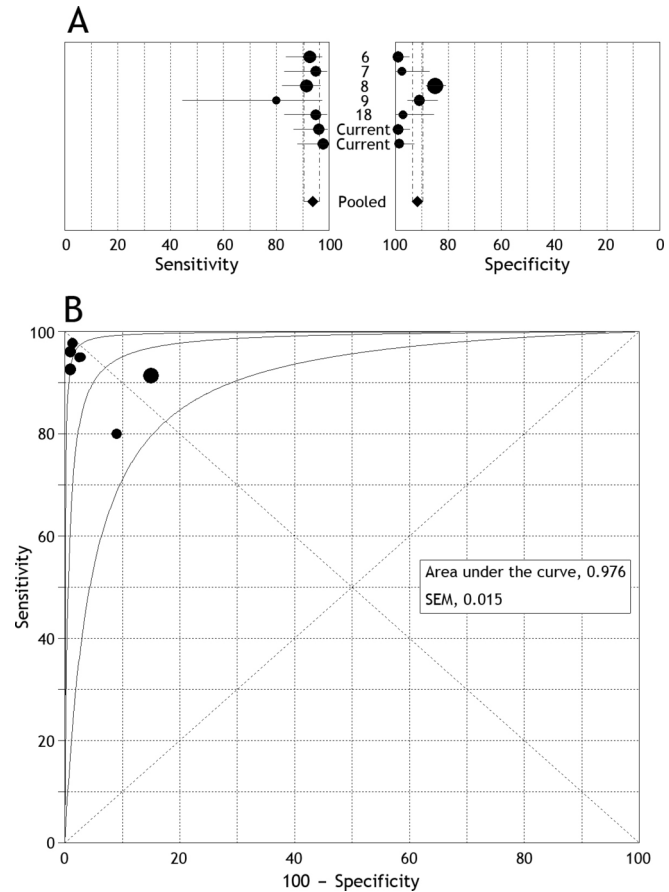

Figure 2 Meta-analysis of overall diagnostic performance of interleukin 27 (IL-27) in differentiating between tuberculous and non-tuberculous effusions. (A) A forest plot of estimates of sensitivity and specificity for IL-27 in pleural fluid for diagnosing tuberculous pleural effusion (TPE). The point estimates of sensitivity and specificity from each study are shown as solid circles. Error bars are 95% confidence intervals. Numbers indicate the reference numbers of studies cited in the reference list. (B) A summary receiver operating characteristic curve with 95% CI for IL-27 in pleural fluid for diagnosing TPE. Each solid circle represents one study in the meta-analysis. The size of each study is indicated by the size of the solid circle, and two solid circles overlapped completely because their sensitivity and specificity were identical.

Where diagnosis of TPE is difficult a number of soluble biomarkers in pleural fluid have been evaluated extensively for this purpose.[27] Two previous meta-analyses showed that nucleic acid amplification tests and T cell IFN-γ release assays cannot replace conventional tests in the diagnosis of TPE.[28 29] Another two meta-analyses showed that pleural ADA and IFN-γ are two valuable biomarkers for such a diagnostic purpose.[4 5] Our study confirmed our previous findings[6 7] and provided stronger evidence to support the use of IL-27 as an excellent diagnostic biomarker for discriminating between TPEs and non-TPEs, with high sensitivity and high specificity of 96.1% and 99.0%, respectively (Beijing cohort).

Our data from the Beijing cohort also showed that a positive likelihood ratio value of 99.0 with IL-27 suggested that the patients with TPE have a 99-fold higher chance of being IL-27-positive than patients without the disease, which is sufficiently high for diagnosis. This may help to avoid unnecessary invasive diagnostic investigations in a number of patients. Moreover, a negative likelihood ratio value of 0 suggests that

Table 5 Post-test probability of tuberculous pleural effusion according to pre-test probability*

Post-test probability	Pre-test probability		
	10%	25%	50%
With positive result (%)	76.7 (45.1 to 92.9)	90.8 (71.2 to 97.5)	96.8 (88.2 to 99.2)
With negative result (%)	0.009 (0.006 to 0.013)	0.025 (0.017 to 0.038)	0.074 (0.048 to 0.107)

*Values in parentheses are 95% CIs estimated by using the lower and upper CIs of the random-effects weighted detection rate and false-positive rate estimates.

if the IL-27 result is negative, there is no probability that this patient has TPE. The high positive predictive value (98.0) and high negative predictive value (98.1) of IL-27 found in this study further indicate that both the false-negative rate and false-positive rates are low. One important strength of this study was that the excellent diagnostic performance of IL-27 found in the Beijing study was verified prospectively in the Wuhan study and further confirmed in the meta-analysis.

Despite the relatively high overall diagnostic accuracy of ADA reported in a previous meta-analysis,[4] intense criticism of its use for diagnosis has been made as its performance is highly dependent on the prevalence of TPE in the studied populations.[30–32] One study from Spain showed that IL-27 is less efficient than ADA in the diagnosis of TPE.[8] In contrast, by a head-to-head comparison of the diagnostic measurements, we found in both the Beijing and Wuhan cohorts that IL-27 is significantly better than ADA for discriminating between TPEs and non-TPEs. In addition, it has been shown that the product of IL-27 and ADA is more accurate than ADA alone in European countries.[8 9] We failed to observe similar results, which is unsurprising as IL-27 alone is itself a good diagnostic indicator. When the product of IL-27 and ADA is used to identify patients with TPE, a false-negative rate of 5% is seen.

The aetiological spectrum of pleural effusions is clearly dependent on the population studied, on the aetiological classification used and on the diagnostic methods employed. In developing countries the common causes of pleural effusions in adults are tuberculosis and bacterial infection,[33] whereas in developed countries more than 90% of all pleural effusions are caused by congestive heart failure, malignancy, pneumonia and pulmonary embolism.[34] More recently, we have reported that of 833 patients with undiagnosed exudative pleural effusion who underwent medical thoracoscopy, 333 (40.0%) were eventually diagnosed as having TPE.[35] In contrast, the proportions of TPE are very low in countries with low levels of tuberculosis, such as New Zealand (5.9%),[36] United Kingdom (5.6%),[37] Spain (3.2%),[38] France (2.7%),[39] Denmark (2.1%),[40] and United States (0%).[41] We noted in our study that the diagnostic accuracy of IL-27 was better than that reported in the two European studies.[8 9] The higher diagnostic accuracy of IL-27 seen in areas with a high incidence of tuberculosis is probably due to the increased prevalence of TPE.

Therefore, it is important to consider the pre-test probability of TPE in interpreting IL-27 results for clinical decision-making. Our meta-analysis indicated that when the predicted probability of TPE is low (10%), the overall post-test probability of positive results of IL-27 is high (76.7%), making it unacceptable for 'ruling in' TPE. However, in high prevalence settings a positive test result increases the probability for TPE up to 96.8%, meaning that further diagnostic procedures can be avoided. In all prevalence settings, when the result of IL-27 is negative, the post-test probabilities are sufficiently low (always <0.1%) to reliably rule out the disease, thus warranting the use of a biomarker-based approach.

Our meta-analysis had several limitations. First, a marked deficiency in the included studies is the use of narrow patient spectrums in the control group in the three studies conducted in China and a small proportion of patients with TPE included in the two studies conducted in Europe. These factors might have biased the estimation of test accuracy.[42 43] Second, the number of studies available for inclusion in this meta-analysis was limited, with one study involving only 10 TPE specimens,[9] and such small studies may be vulnerable to selection bias. Third, there might have been misclassification bias. A small proportion of patients with TPE were not diagnosed on the basis of bacteriological or histological results, but just on the clinical course.[6 7 9] This might have caused non-random misclassification, leading to biased results. Fourth, control populations were heterogeneous among the studies; and various cut-off points were used for distinguishing between TPE and non-TPEs according to the best combination of sensitivity and specificity. These problems with the accuracy of diagnosis might also lead to biased results.

It should be mentioned that because the determination of ADA is simpler and less expensive, the current guideline recommends it as the preferred test for the diagnosis of TPE.[32] In contrast, despite the superiority of IFN-γ in diagnosing TPE, the high cost and the lack of an accepted cut-off value for the discrimination of TPE limit its daily clinical use.[2 27] Based on the excellent diagnostic performance of IL-27 for TPE found in this study, it is reasonable to recommend the application of pleural IL-27 measurement in routine clinical practice.

In conclusion, our prospective study findings have provided sufficient evidence to support the use of IL-27 as a valuable indicator for the diagnosis of TPE, with a diagnostic performance comparable to that of IFN-γ and significantly more accurate than that of ADA. Our meta-analysis data further support the use of IL-27 to rule in TPE in an area with high prevalence, and can also be reliably used as a 'rule out' test in all prevalence settings. It should be noted that since disparate cut-off values exist in different studies included in the meta-analysis, more extensive studies are needed to confirm these results, especially to establish an acceptable cut-off point for clinical practice.

Competing interests None declared.

Patient consent Obtained.

Ethics approval Institutional review board for human studies of Beijing Chaoyang Hospital; and institutional review board for human studies of Wuhan Union Hospital, China.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Porcel JM. Tuberculous pleural effusion. *Lung* 2009;187:263–70.
2 Light RW. Update on tuberculous pleural effusion. *Respirology* 2010;15:451–8.
3 Gopi A, Madhavan SM, Sharma SK, *et al*. Diagnosis and treatment of tuberculous pleural effusion in 2006. *Chest* 2007;131:880–9.
4 Liang QL, Shi HZ, Wang K, *et al*. Diagnostic accuracy of adenosine deaminase in tuberculous pleurisy: a meta-analysis. *Respir Med* 2008;102:744–54.
5 Jiang J, Shi HZ, Liang QL, *et al*. Diagnostic value of interferon-gamma in tuberculous pleurisy: a metaanalysis. *Chest* 2007;131:1133–41.
6 Yang WB, Liang QL, Ye ZJ, *et al*. Cell origins and diagnostic accuracy of interleukin 27 in pleural effusions. *PLoS One* 2012;7:e40450.
7 Wu YB, Ye ZJ, Qin SM, *et al*. Combined detections of interleukin 27, interferon-γ, and adenosine deaminase in pleural effusion for diagnosis of tuberculous pleurisy. *Chin Med J* 2013;126:3215–21.
8 Valdés L, San José E, Ferreiro L, *et al*. Interleukin 27 could be useful in the diagnosis of tuberculous pleural effusions. *Respir Care* 2014;59:399–405.
9 Skouras VS, Magkouta SF, Psallidas I, *et al*. Interleukin-27 improves the ability of adenosine deaminase to rule out tuberculous pleural effusion regardless of pleural tuberculosis prevalence. *Infect Dis* 2015;47:477–83.
10 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29–36.
11 Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem* 1993;39:561–77.
12 Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983;148:839–43.
13 Devillé WL, Buntinx F, Bouter LM, *et al*. Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol* 2002;2:9.
14 Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat Med* 1993;12:1293–316.
15 Irwig L, Macaskill P, Glasziou P, *et al*. Meta-analytic methods for diagnostic test accuracy. *J Clin Epidemiol* 1995;48:119–30.
16 Vamvakas EC. Meta-analyses of studies of the diagnostic accuracy of laboratory tests: a review of the concepts and methods. *Arch Pathol Lab Med* 1998;122:675–86.
17 Egger M, Davey Smith G, Schneider M, *et al*. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629–34.
18 Sun M, Yan D, Jiang S, *et al*. [Diagnostic value of interleukin-27 in tuberculous pleural effusion]. *Zhonghua Yi Xue Za Zhi* 2014;94:2641–4.
19 Whiting P, Rutjes AW, Reitsma JB, *et al*. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
20 Bossuyt PM, Reitsma JB, Bruns DE, *et al*. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 2003;326:41–4.
21 Diacon AH, Van de Wal BW, Wyser C, *et al*. Diagnostic tools in tuberculous pleurisy: a direct comparative study. *Eur Respir J* 2003;22:589–91.
22 Escudero Bueno C, García Clemente M, Cuesta Castro B, *et al*. Cytologic and bacteriologic analysis of fluid and pleural biopsy specimens with Cope's needle. Study of 414 patients. *Arch Intern Med* 1990;150:1190–4.
23 Valdés L, Alvarez D, San José E, *et al*. Tuberculous pleurisy: a study of 254 patients. *Arch Intern Med* 1998;158:2017–21.
24 Epstein DM, Kline LR, Albelda SM, *et al*. Tuberculous pleural effusions. *Chest* 1987;91:106–9.
25 Seibert AF, Haynes J, Middleton R, *et al*. Tuberculous pleural effusion. Twenty-year experience. *Chest* 1991;99:883–6.
26 Wang Z, Xu LL, Wu YB, *et al*. Diagnostic value and safety of medical thoracoscopy in tuberculous pleural effusion. *Respir Med* 2015;109:1188–92.
27 Skouras VS, Kalomenidis I. Pleural fluid tests to diagnose tuberculous pleuritis. *Curr Opin Pulm Med* 2016;22:367–77.
28 Pai M, Flores LL, Hubbard A, *et al*. Nucleic acid amplification tests in the diagnosis of tuberculous pleuritis: a systematic review and meta-analysis. *BMC Infect Dis* 2004;4:6.
29 Zhou Q, Chen YQ, Qin SM, *et al*. Diagnostic accuracy of T-cell interferon-γ release assays in tuberculous pleurisy: a meta-analysis. *Respirology* 2011;16:473–80.
30 Valdés L, San José ME, Pose A, *et al*. Diagnosing tuberculous pleural effusion using clinical data and pleural fluid analysis. A study of patients less than 40 years-old in an area with a high incidence of tuberculosis. *Respir Med* 2010;104:1211–7.
31 Garcia-Zamalloa A, Taboada-Gomez J. Diagnostic accuracy of adenosine deaminase and lymphocyte proportion in pleural fluid for tuberculous pleurisy in different prevalence scenarios. *PLoS One* 2012;7:e38729.
32 Hooper C, Lee YC, Maskell N, *et al*. Investigation of a unilateral pleural effusion in adults: British Thoracic Society Pleural Disease Guideline 2010. *Thorax* 2010;65 (Suppl 2):ii4–17.
33 Afful B, Murphy S, Antunes G, *et al*. The characteristics and causes of pleural effusions in Kumasi Ghana - a prospective study. *Trop Doct* 2008;38:219–20.
34 Light RW. Clinical practice. Pleural effusion. *N Engl J Med* 2002;346:1971–7.
35 Wang XJ, Yang Y, Wang Z, *et al*. Efficacy and safety of diagnostic thoracoscopy in undiagnosed pleural effusions. *Respiration* 2015;90:251–5.
36 Wilsher ML, Veale AG. Medical thoracoscopy in the diagnosis of unexplained pleural effusion. *Respirology* 1998;3:77–80.
37 Munavvar M, Khan MA, Edwards J, *et al*. The autoclavable semirigid thoracoscope: the way forward in pleural disease? *Eur Respir J* 2007;29:571–4.
38 Ferrer J, Roldán J, Teixidor J, *et al*. Predictors of pleural malignancy in patients with pleural effusion undergoing thoracoscopy. *Chest* 2005;127:1017–22.
39 Blanc FX, Atassi K, Bignon J, *et al*. Diagnostic value of medical thoracoscopy in pleural disease: a 6-year retrospective study. *Chest* 2002;121:1677–83.
40 Hansen M, Faurschou P, Clementsen P. Medical thoracoscopy, results and complications in 146 patients: a retrospective study. *Respir Med* 1998;92:228–32.
41 DePew ZS, Wigle D, Mullon JJ, *et al*. Feasibility and safety of outpatient medical thoracoscopy at a large tertiary medical center: a collaborative medical-surgical initiative. *Chest* 2014;146:398–405.
42 Lijmer JG, Mol BW, Heisterkamp S, *et al*. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;282:1061–6.
43 Leeflang MM, Bossuyt PM, Irwig L. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. *J Clin Epidemiol* 2009;62:5–12.