Diagnostic accuracy of tumour markers for malignant pleural effusion: a meta-analysis

Q-L Liang, H-Z Shi, X-J Qin, X-D Liang, J Jiang, H-B Yang

▶ Additional data are published in the online Appendix at http://thorax.bmj.com/content/vol63/issue1

Institute of Respiratory Diseases, First Affiliated Hospital, Guangxi Medical University, Nanning, Guangxi, People's Republic of China

Correspondence to:
Dr H-Z Shi, Institute of Respiratory Diseases, First Affiliated Hospital, Guangxi Medical University, Nanning 530021, Guangxi, People's Republic of China; shihuanzhong@sina.com

ABSTRACT

Background: The role of tumour markers such as carbohydrate antigen (CA) 125, CA 15-3, CA 19-9 and CYFRA 21-1 (a fragment of cytokeratin 19) in differentiating malignant pleural effusions (MPE) from benign effusions is not yet clear.

Methods: After a systematic review of English language studies, sensitivity, specificity and other measures of accuracy of pleural concentrations of CA 125, CA 15-3, CA 19-9 and CYFRA 21-1 or their combinations in the diagnosis of MPE were pooled using random effects models. Summary receiver operating characteristic curves were used to summarise overall test performance.

Results: Twenty-nine studies met the inclusion criteria for the analysis. The summary estimates of the sensitivity and specificity of these tumour markers were as follows: CA 125, 0.48/0.85; CA 15-3, 0.51/0.96; CA 19-9, 0.25/0.96; CYFRA 21-1, 0.55/0.91 for diagnosing MPE. The estimated summary receiver operating characteristic curves showed that the performance of pleural CA 125 and CA 19-9 measurement in the diagnosis of MPE was limited, whereas that of CA 15-3 and CYFRA 21-1 was better. When two or more of the above four tumour markers were combined, or combined with carcinoembryonic antigen, the sensitivity and specificity were all increased to different extents.

Conclusions: The current evidence does not recommend using one tumour marker alone for the diagnosis of MPE, but the combination of two or more tumour markers seems to be more sensitive. The results of tumour marker assays should be interpreted in parallel with clinical findings and the results of conventional tests.

Pleural effusions may occur in patients suffering from physical trauma or systemic disorders such as infection, inflammation or cancer.[1] Malignancy is one of the main causes of pleural effusions, and >90% of malignant pleural effusions (MPE) are due to metastatic disease.[2] Carcinoma of any organ can metastasise to the pleura, but the most frequent are lung and breast carcinomas and lymphomas, less frequently digestive and ovary carcinomas.[3] It is important to elucidate their precise aetiologies to differentiate MPE from benign effusions. Cytological examination is a standard method for the diagnosis of MPE. Although repeated thoracenteses can increase the sensitivity of cytology, it is typically only 50–70%.[2] Blind pleural biopsy can be performed in addition. However, among cytology-negative cases, only 7–13% are proved to be MPE by an additional biopsy.[4 5]

To improve the diagnosis of MPE, a number of tumour markers have been intensively evaluated but the search for a highly accurate tumour marker

in pleural fluid that reliably confirms MPE has so far been fruitless.[6] Carcinoembryonic antigen (CEA) is the most common marker to have been studied extensively and has been found to be of diagnostic significance. Carbohydrate antigen (CA) 125 is a tumour-associated antigen commonly seen in ovarian carcinoma and is used to assess the response to chemotherapy and for early detection of relapse.[7] Using immunostaining of cells in pleural fluid specimens with anti-CA 15-3 antibody, the sensitivity of CA 15-3 was 91% for breast carcinoma and 80% for all adenocarcinomas and the specificity was 94% for breast carcinoma and for all adenocarcinomas.[8] CA 19-9 is a tumour antigen whose level increases particularly in gastrointestinal tumours. Molina and coworkers[9] reported a high level of CA 19-9 in the serum of patients with lung cancer. CYFRA 21-1 is a fragment of cytokeratin 19 which provides a useful marker for epithelial malignancies, distinctly reflecting ongoing cell activity.[10] Although the diagnostic accuracy of CA 125, CA 15-3, CA 19-9 or/and CYFRA 21-1 for MPE has been extensively studied, the exact roles of these detections remain controversial. We performed a meta-analysis to establish the overall diagnostic accuracy of pleural CA 125, CA 15-3, CA 19-9, CYFRA 21-1 and CEA (either singly or in combination) for MPE.

METHODS

Search strategy and study selection

Embase, Ovid, Web of Science, the Cochrane database and Medline (using PubMed as the search engine) were searched to identify suitable studies up to 30 December 2006; no start date limit was applied. Articles were also identified by use of the related articles function in PubMed. References of articles identified were also searched manually. The search terms were "tumour marker", "carbohydrate antigen 125/CA 125", "carbohydrate antigen 15-3/CA 15-3", "carbohydrate antigen 19-9/CA 19-9", "fragment of cytokeratin 19/CYFRA 21-1", "lung cancer", "malignant pleural mesothelioma", "pleural effusion/pleural fluid", "sensitivity and specificity" and "accuracy". Although no language restrictions were imposed initially, for the full-text review and final analysis our resources only permitted review of English articles. Conference abstracts and letters to journal editors were excluded because of the limited data presented in them.

A study was included in the meta-analysis if it provided CA 125, CA 15-3, CA 19-9 and CYFRA 21-1 values for both sensitivity and specificity of the diagnosis of MPE. Only studies including at least 10 pleural fluid specimens were selected since

very small studies may be vulnerable to selection bias. Publications with evidence of a possible overlap of patients with other studies were discussed by QLL, HZS and XJQ and only the best quality study was used. Two reviewers (QLL and HZS) independently judged study eligibility while screening the citations. Disagreements were resolved by consensus.

Data extraction and quality assessment

The final set of English articles was assessed independently by two reviewers (QLL and HZS). The reviewers were blinded to publication details and disagreements were resolved by consensus. Data retrieved from the reports included author, publication year, participant characteristics, test methods, sensitivity and specificity data, cut-off value and methodological quality.

The methodological quality of the studies was assessed using guidelines published by the STARD (standards for reporting diagnostic accuracy, maximum score 25) initiative[11] (ie, guidelines that aim to improve the quality of reporting in diagnostic studies) and the QUADAS (quality assessment for studies of diagnostic accuracy, maximum score 14) tool[12] (ie, appraisal by use of empirical evidence, expert opinion and formal consensus to assess the quality of primary studies of diagnostic accuracy). In addition, for each study the following characteristics of study design were retrieved: (1) cross-sectional design (versus case-control design); (2) consecutive or random sampling of patients; (3) blinded (single or double) interpretation of determination and reference standard results; (4) prospective data collection. If no data on the above criteria were reported in the primary studies, the information was requested from the authors. If the authors did not respond, the "unknown" items were treated as "No".

Statistical analyses

Standard methods recommended for meta-analyses of diagnostic test evaluations were used.[13] Analyses were performed using the following statistical software programs: Stata Version 8.2 (Stata Corporation, College Station, Texas, USA); Meta-Test Version 0.6 (New England Medical Center, Boston, Massachusetts, USA) and Meta-DiSc for Windows (XI Cochrane Colloquium; Barcelona, Spain). The following measures of test accuracy were computed for each study: sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR) and diagnostic odds ratio (DOR).

The analysis was based on a summary receiver operating characteristic (SROC) curve.[13 14] The sensitivity and specificity for the single test threshold identified for each study were used to plot an SROC curve[14 15] (see statistical methods given in the Appendix available online only). A random-effects model was used to calculate the average sensitivity, specificity and the other measures across studies.[16 17]

The χ^2 and Fisher exact tests were used to detect statistically significant heterogeneity across the studies. To assess the effects of STARD and QUADAS scores on the diagnostic ability of CA 125, CA 15-3, CA 19-9, and CYFRA 21-1, we included them and the study design characteristics as covariates in univariate meta-regression analysis (inverse variance weighted). The relative DOR (RDOR) was calculated according to standard methods to analyse the change in diagnostic precision in the study per unit increase in the covariate.[18 19] Since publication bias is of concern for meta-analyses of diagnostic studies, we tested for the potential presence of this bias using funnel plots and the Egger test.[20]

RESULTS

After independent review, 59 non-English publications were excluded from the meta-analysis (publication list available on request); 46 publications dealing with pleural concentrations of CA 125, CA 15-3, CA 19-9 and/or CYFRA 21-1 for diagnosis of MPE were considered to be eligible for inclusion in the analysis (Appendix references A1–A46 available in the online supplement). Of these publications, two were excluded because they recruited fewer than 10 patients in one of the study groups (references A30 and A31), three were excluded because the tumour marker concentrations were determined only in cases of MPE (references A32–A34), eight were excluded because they included MPE and malignant peritoneal effusions as a single group (references A35–A42), one was excluded because it did not allow the calculation of sensitivity or specificity (reference A43), and three were excluded because the same authors published several reports on the same patients and only the best quality study was considered (references A44–A46). A total of 29 articles (A1–A29) were therefore available for analysis of the diagnosis accuracy of CA 125, CA 15-3, CA 19-9 and/or CYFRA 21-1 in MPE.

Quality of reporting and study characteristics

As shown in Appendix table 1 (available in online supplement), 9 of the 29 studies (31.0%) had a cross-sectional design; in 17 studies (58.6%) the samples were collected from consecutive patients; 14 studies (48.3%) reported blinded interpretation of tumour marker assays independent of the reference standard; and 13 studies (44.8%) had a prospective study design. The clinical characteristics, together with STARD and QUADAS scores of studies of CA 125, CA 15-3, CA 19-9 and CYFRA 21-1, are outlined in Appendix tables 2, 3, 4 and 5, respectively (available in the online supplement). The average sample size of the CA 125, CA 15-3, CA 19-9 and CYFRA 21-1 studies was 131 (range 41–416), 162 (range 39–416), 138 (range 61–336) and 127 (range 35–416), respectively. In the present meta-analysis, all patients with pleural malignancies were confirmed based on the conventional "gold standard"; MPE was confirmed by cytological study, pleural biopsy specimens or necropsy.

Diagnostic accuracy

Figure 1 shows the forest plots of sensitivities and specificities for pleural concentrations of the four tumour markers in the diagnoses of MPE. The graphs of the SROC curves for the determinations of tumour markers showing true positive rates versus false positive rates from individual studies are indicated in fig 2. Pooled results of the diagnostic accuracy of each tumour marker in MPE are shown in table 1.

The sensitivity of pleural CA-125 measurement in the diagnosis of MPE varied between 0.17 and 1.00, and the specificity varied between 0.05 and 1.00; PLR was 5.96 (range 1.06–85.67), NLR was 0.54 (range 0.02–0.83) and DOR was 19.61 (range 2.14–731.00). For CA 15-3, the sensitivity ranged from 0.30 to 0.80, while specificity ranged from 0.75 to 1.00; PLR was 11.69 (range 2.05–151.80), NLR was 0.52 (range 0.22–0.70) and DOR was 24.74 (range 3.15–224.00). For CA 19-1, the sensitivity ranged from 0.13 to 0.89, while specificity ranged from 0.73 to 1.00; PLR was 10.42 (range 1.42–58.50), NLR was 0.70 (range 0.11–0.87) and DOR was 19.88 (range 1.69–348.50). For CYFRA 21-1, the sensitivity ranged from 0.20 to 0.91, while specificity ranged from 0.08 to 1.00; PLR was 6.55 (range 0.99–103.43), NLR was 0.43 (range 0.10–0.1.18) and DOR was 16.24 (range 0.83–130.02).
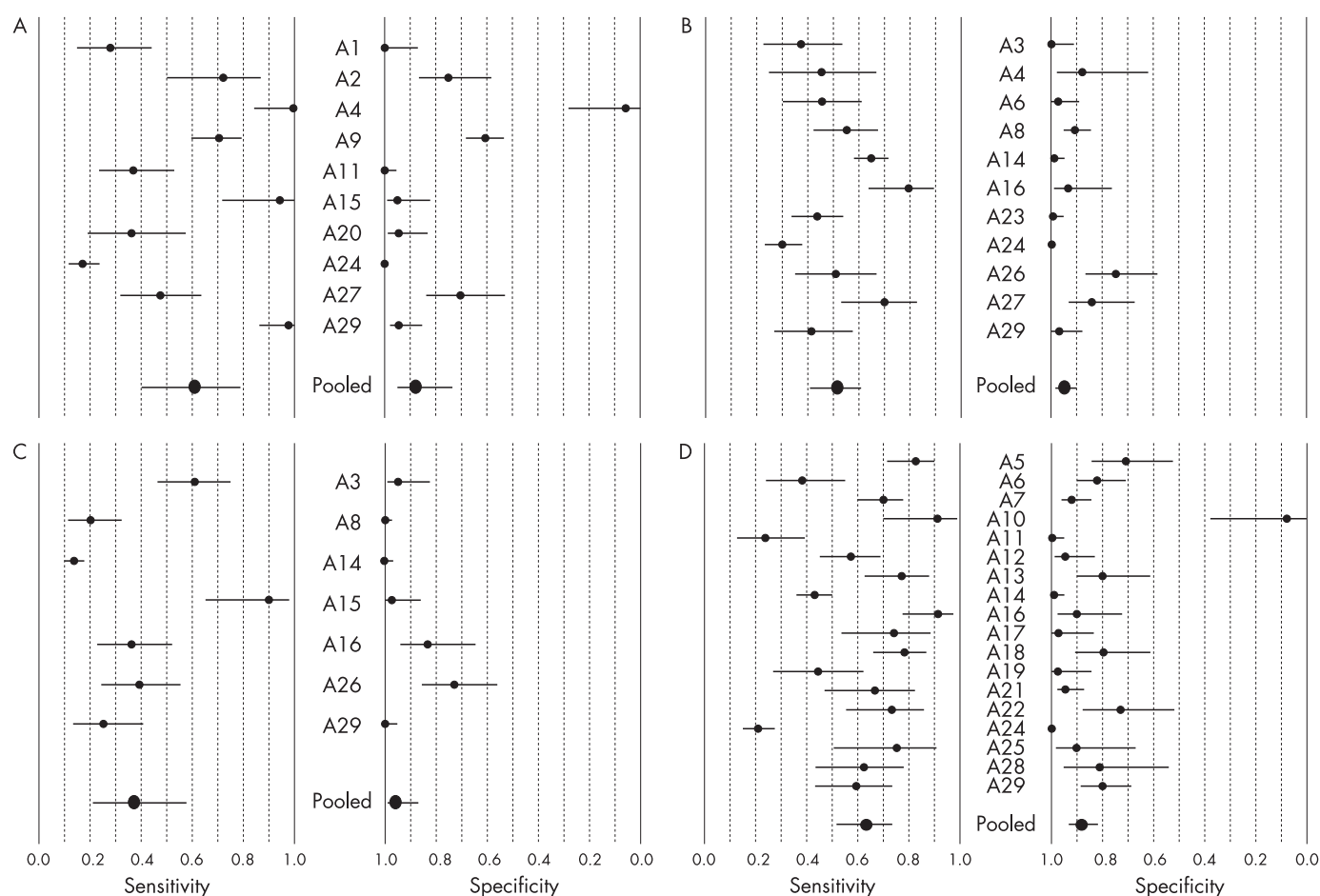
Figure 1 Forest plots of estimates of sensitivity and specificity for assays of (A) CA 125, (B) CA 15-3, (C) CA 19-9 and (D) CYFRA 21-1 in the diagnosis of malignant pleural effusions. The point estimates of sensitivity and specificity from each study are shown as solid circles. Error bars are 95% confidence intervals. Numbers indicate the reference numbers of studies cited in the reference list in the Appendix (available online only).

As shown in table 1, the Q values for sensitivity, specificity, PLR, NLR and DOR of studies of all the tumour markers were high, with all p values <0.001 indicating significant heterogeneity between all studies.

The SROC curve and its area under the curve (AUC) present an overall summary of test performance and display the trade-off between sensitivity and specificity. In all 29 studies included in the meta-analysis, both sensitivity and specificity were indicated directly in each publication and a single optimal cut-off value was selected by all investigators and was reported in their publications, respectively. The meta-analysis showed that the mean values of the maximum joint sensitivity and specificity of CA 125, CA 15-3, CA 19-9 and CYFRA 21-1 were 0.81, 0.68, 0.72 and 0.76, respectively, and their mean AUCs were 0.88, 0.73, 0.78 and 0.83, respectively, indicating that the overall accuracy was not as high as expected.

In the publications included in the meta-analysis, some studies evaluated the simultaneous determination of two or more pleural tumour markers in the diagnosis of MPE. The pooled results of the diagnostic accuracy of a combination of two or more tumour markers of CA 125, CA 15-3, CA 19-9, CYFRA 21-1 and CEA are shown in table 2. The results indicate that some combinations of tumour markers have a greater diagnostic role than one tumour marker alone.

Multiple regression analysis and publication bias

The scores of both STARD and QUADAS were used in the meta-regression analysis to assess the effect of study quality on RDOR of CA 125, CA 15-3, CA 19-9 and CYFRA 21-1 in the diagnosis of MPE. Table 3 shows the findings of the meta-regression analysis of the RDOR (dependent variable) between studies with higher and lower study quality scores. CA 125 assays of higher quality (STARD ≥13) produced RDOR that were not significantly higher than studies of lower quality (STARD <13). Also, studies with QUADAS ≥10 did not have a better performance than those with QUADAS <10. Similarly, studies of CA 15-3, CA19-9 and CYFRA 21-1 of higher quality did not have a better test performance than those of lower quality. Differences in studies with or without blind design, cross-sectional, consecutive/random and prospective design did not reach statistical significance (data not shown). These results indicate that the study quality and design did not substantially affect the accuracy of pleural tumour markers in the diagnosis of MPE.

Evaluation of publication bias showed that the Egger tests for studies of CA 125 (p = 0.014), CA 15-3 (p = 0.015), CA 19-9 (p = 0.009) and CYFRA 21-1 (p = 0.043) in the diagnosis of MPE were all significant. Four funnel plots for publication bias show some asymmetry (see fig 1 in Appendix available online only). These results indicate a potential for publication bias.

Figure 2 Summary receiver operating characteristic curves for assays of (A) CA 125, (B) CA 15-3, (C) CA 19-9 and (D) CYFRA 21-1. Solid circles represent individual studies in the meta-analysis. The size of each study is indicated by the size of the solid circle. The weighted (solid line) and unweighted (broken line) regression summary receiver operating characteristic curves summarise the overall diagnostic accuracy.
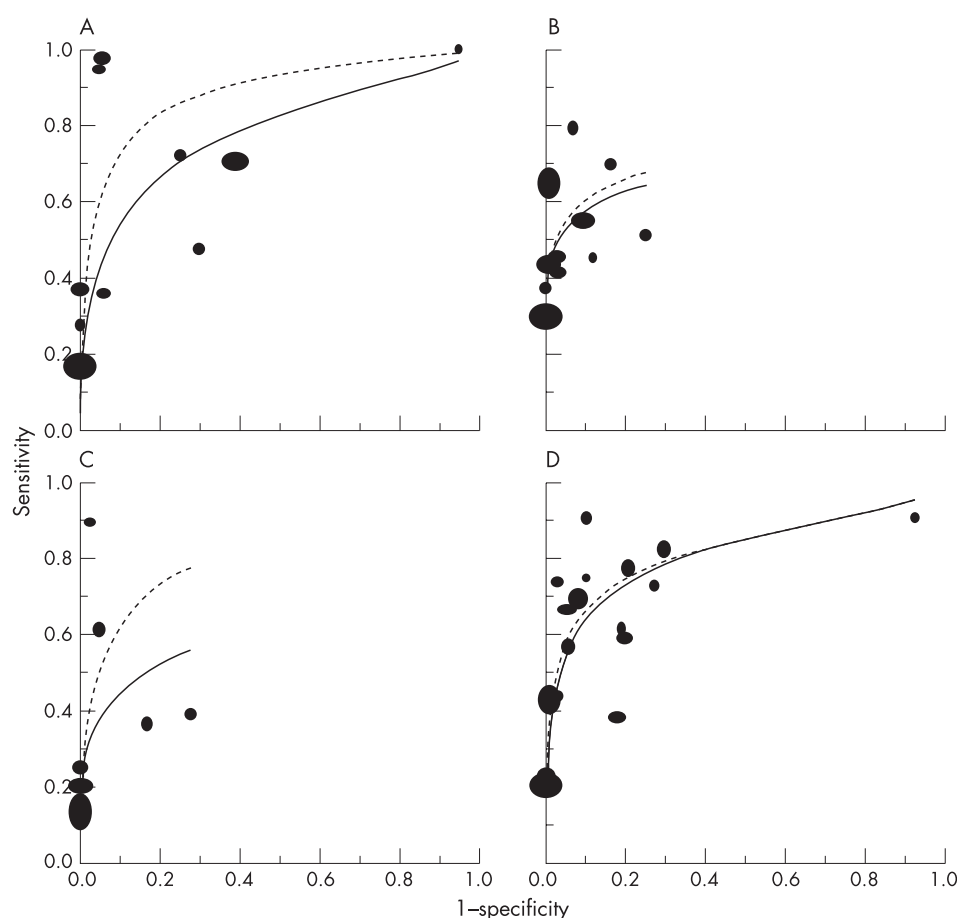
DISCUSSION

Making a differential diagnosis between MPE and non-MPE is a critical clinical problem and conventional tests are not always.[21] Determination of tumour markers in pleural fluid has been proposed as an alternative non-invasive way of establishing a diagnosis of MPE.[6]

The overall specificity of CA 15-3, CA 19-9 and CYFRA 21-1, but not CA 125, was more than 0.90. The summary estimate of the sensitivities for the four tumour markers were, however, all quite low and were more variable than the specificity. These data suggest a potential role for determination of these tumour

markers in confirming (ruling in) MPE. However, these tests maximise specificity at the cost of sensitivity, and this trade-off has significant clinical implications. By contrast with the higher specificity, these tumour markers had low sensitivities that were not sufficiently low to exclude non-MPE when the pleural tumour marker concentrations are lower than the cut-off values. Negative tests do not therefore mean absence of MPE, and patients with negative tumour marker results have a fairly high chance of having MPE.

The SROC curve presents a global summary of test performance and shows the trade-off between sensitivity and

Table 1 Pooled results of diagnostic accuracy of each tumour marker in malignant pleural effusions (MPE)

	CA 125	CA 15-3	CA 19-9	CYFRA 21-1
Number of studies	10	11	7	18
Number of patients with MPE/non-MPE	512/801	819/966	598/488	1152/1122
Sensitivity (95% CI)	0.48 (0.44 to 0.53)	0.51 (0.47 to 0.54)	0.25 (0.21 to 0.28)	0.55 (0.52 to 0.58)
Heterogeneity* (p)	212.65 (<0.001)	76.46 (<0.001)	99.32 (<0.001)	246.24 (<0.001)
Specificity (95% CI)	0.85 (0.83 to 0.88)	0.96 (0.95 to 0.97)	0.96 (0.94 to 0.98)	0.91 (0.90 to 0.93)
Heterogeneity (p)	260.86 (<0.001)	72.58 (<0.001)	60.97 (<0.001)	177.54 (<0.001)
PLR (95% CI)	5.96 (2.27 to 15.68)	11.69 (5.05 to 27.07)	10.42 (2.66 to 40.78)	6.55 (3.09 to 13.88)
Heterogeneity (p)	311.44 (<0.001)	55.65 (<0.001)	36.90 (<0.001)	313.76 (<0.001)
NLR (95% CI)	0.54 (0.38 to 0.78)	0.52 (0.43 to 0.62)	0.70 (0.55 to 0.88)	0.43 (0.34 to 0.54)
Heterogeneity (p)	109.99 (<0.001)	63.21 (<0.001)	69.30 (<0.001)	198.87 (<0.001)
DOR (95% CI)	19.61 (6.25 to 61.52)	24.74 (10.50 to 58.28)	19.88 (4.19 to 94.24)	16.24 (9.60 to 27.49)
Heterogeneity (p)	49.68 (<0.001)	36.96 (<0.001)	32.75 (<0.001)	47.23 (<0.001)
AUC (SEM)	0.88 (0.06)	0.73 (0.07)	0.78 (0.16)	0.83 (0.03)

*Q value.

CI, confidence interval; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio; AUC, area under curve; SEM, standard error of mean.

Table 2 Pooled results of diagnostic accuracy of tumour marker combinations in malignant pleural effusions (MPE)

Combinations	Number of studies	Number of patients with MPE/non-MPE	Sensitivity (95% CI)	Specificity (95% CI)	PLR (95% CI)	NLR (95% CI)	DOR (95% CI)	AUC (SEM)	Reference*
CEA + CA 125	3	107/167	0.65 (0.56 to 0.74)	0.98 (0.94 to 0.99)	25.20 (4.07 to 156.07)	0.35 (0.21 to 0.57)	66.01 (16.13 to 269.40)	0.86 (0.10)	A1, A11, A20
CEA + CA 15-3	3	208/366	0.66 (0.59 to 0.72)	0.98 (0.96 to 0.99)	26.37 (12.05 to 57.72)	0.35 (0.29 to 0.43)	86.1 (39.0 to 190.2)	0.88 (0.12)	A3, A8, A23
CEA + CA 19-9	1	65/142	0.46	1.00	132.17 (8.21 to 2128.70)	0.54 (0.43 to 0.68)	244.86 (14.62 to 41.0.90)	NA	A8
CEA + CYFRA 21-1	3	97/115	0.72 (0.62 to 0.81)	0.95 (0.89 to 0.98)	8.89 (1.99 to 39.69)	0.24 (0.08 to 0.68)	48.48 (15.52 to 151.42)	0.93 (0.03)	A11, A25, A28
CA 125 + CA 19-9	1	19/42	0.95	0.98	38.79 (5.72 to 276.72)	0.05 (0.01 to 0.36)	738.00 (43.70 to 12464.20)	NA	A15
CA 125 + CYFRA 21-1	2	115/133	0.66 (0.57 to 0.75)	0.96 (0.90 to 0.98)	17.9 (1.2 to 276.3)	0.38 (0.28 to 0.54)	45.0 (4.8 to 422.0)	NA	A11, A12
CA 15-3 + CYFRA 21-1	1	65/142	0.62	0.97	21.85 (8.16 to 58.50)	0.40 (0.29 to 0.54)	55.20 (18.14 to 167.93)	NA	A8
CA 15-3 + CYFRA 21-1	2	87/109	0.83 (0.73 to 0.90)	0.95 (0.90 to 0.99)	19.03 (0.71 to 510.81)	0.08 (0.00 to 4.35)	345.53 (44.74 to 2668.80)	NA	A11, A15
CEA + CA 125 + CYFRA 21-1	1	43/79	0.65	1.00	103.64 (6.48 to 1656.70)	0.35 (0.24 to 0.53)	292.35 (16.94 to 5046.60)	NA	A11
CEA + CA 15-3 + CA 19-9	1	65/142	0.71	0.97	25.12 (9.44 to 66.85)	0.30 (0.21 to 0.44)	83.53 (27.02 to 258.23)	NA	A8

*Appendix references are available in online data supplement.
CI, confidence interval; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio; AUC, area under curve; SEM, standard error of mean; NA, not applicable.

specificity. As a global measure of test efficacy, we used the maximum joint sensitivity and specificity—the point of intersection of the SROC curve with a diagonal line from the left upper corner to the right lower corner of the SROC space—which corresponds to the highest common value of sensitivity and specificity for the test.[14] This point does not indicate the only (or even the best) combination of sensitivity and specificity for a particular clinical setting, but it represents an overall measure of the discriminatory power of a test. Our data showed that the values of the maximum joint sensitivity and specificity of CA 125, CA 15-3, CA 19-9 and CYFRA 21-1 were 0.81, 0.68, 0.72 and 0.76, respectively. On the other hand, their mean AUCs were 0.88, 0.73, 0.78 and 0.83, respectively. All these data suggest that the overall accuracy of tumour markers in diagnosing MPE are not as high as expected.

The DOR is a single indicator of test accuracy[22] that combines the data from sensitivity and specificity into a single number. The DOR of a test is the ratio of the odds of a positive test result in a subject with the disease relative to the odds of a positive test result in a subject without the disease. The value of a DOR ranges from 0 to infinity, with higher values indicating better discriminatory test performance (higher accuracy). A DOR of 1.0 indicates that a test does not discriminate between patients with the disorder and those without it. DOR values <1.00 suggest improper test interpretation (a greater proportion of negative test results in the group with disease). In the present meta-analysis, we found that the mean DOR values for CA 125, CA 15-3, CA 19-9 and CYFRA 21-1 were 19.61, 24.74, 19.88 and 16.24, respectively, indicating that, although not as good as expected, measurement of these four tumour markers could be helpful in the diagnosis of MPE.

Since the SROC curve and the DOR are not easy to interpret and use in clinical practice,[23] and since the likelihood ratios are considered more clinically meaningful,[23 24] we also presented both PLR and NLR as our measures of diagnostic accuracy for the tumour markers. Likelihood ratios of >10 or <0.1 generate large and often conclusive shifts from pre-test to post-test probability (indicating high accuracy).[24] Out data showed that overall PLR values of CA 125, CA 15-3, CA 19-9 and CYFRA 21-1 were 5.96, 11.69, 10.42 and 6.55, respectively, suggesting that patients with MPE have a near 6-fold higher chance of being CA 125 test-positive, a near 12-fold higher chance of being CA 15-3 test-positive, about a 10-fold higher chance of being CA 19-9 test-positive, and a near 7-fold higher chance of being CYFRA test-positive, respectively, compared with patients without MPE. On the other hand, the mean NLR values of CA 125, CA 15-3, CA 19-9 and CYFRA 21-1 were 0.54, 0.52, 0.70 and 0.43, respectively, so if the assay results of CA 125, CA 15-3, CA 19-9 and CYFRA 21-1 were negative, the probability that this patient has MPE is about 54%, 52%, 70% and 43%, respectively, which is too high to rule out MPE.

In addition to the four tumour markers analysed in the present meta-analysis, other biomarkers such as CEA,[25] neuron-specific enolase,[26] CA 549 and CA72-4 have been evaluated for their use in the diagnosis of MPE. In a recent meta-analysis we found that the summary estimates for CEA in the diagnosis of MPE were: sensitivity 0.54 (95% CI 0.52 to 0.55), specificity 0.94 (95% CI 0.93 to 0.95), PLR 9.52 (95% CI 6.97 to 13.01), NLR 0.49 (95% CI 0.44 to 0.54) and DOR 22.5 (95% CI 15.6 to 32.5) (unpublished results). In the present meta-analysis we found that the combination of two or more of CA 125, CA 15-3, CA 19-9 and CYFRA 21-1, as well as CEA, resulted in a greater sensitivity than that of any one of these tumour markers alone.

Table 3 Weighted meta-regression of the effects of methodological quality on diagnostic precision of pleural tumour markers

	Number of studies	Coefficient	RDOR (95% CI)	p Value
CA 125 studies	10			
STARD ≥13	5	1.855	6.39 (0.01 to 2909.97)	0.487
QUADAS ≥10	6	−1.715	0.18 (0.00 to 70.15)	0.508
CA 13-5 studies	11			
STARD ≥13	8	1.857	6.40 (0.77 to 52.92)	0.076
QUADAS ≥10	8	−0.783	0.46 (0.06 to 3.50)	0.393
CA 19-9 studies	7			
STARD ≥13	4	−2.402	0.09 (0.00 to 556.75)	0.445
QUADAS ≥10	4	0.397	1.49 (0.00 to 720.09)	0.851
CYFRA 21-1 studies	18			
STARD ≥13	9	−0.262	0.77 (0.23 to 2.59)	0.650
QUADAS ≥10	10	0.075	1.08 (0.33 to 3.55)	0.894

RDOR, relative diagnostic odds ratio; CI, confidence interval; STARD, standards for reporting diagnostic accuracy; QUADAS, quality assessment for studies of diagnostic accuracy.

An exploration of the reasons for heterogeneity rather than computation of a single summary measure is an important goal of meta-analysis.[27] The regression coefficients for the variables give a measure of the difference in diagnostic accuracy of tumour markers in the two groups, with positive coefficients indicating better discriminating power and negative coefficients corresponding to reduced discriminatory ability. In our meta-analysis, for all four tumour makers analysed, both STARD and QUADAS scores were used in the meta-regression analysis to assess the effect of study quality on RDOR. We did not find that studies of higher quality had a better test performance than those of lower quality, although we found a significant heterogeneity for sensitivity, specificity, PLR, NLR and DOR between these studies. We also noted that differences for studies with or without blinded design, cross-sectional, consecutive/random and prospective design did not reach statistical significance.

Our meta-analysis had some limitations. First, the exclusion of conference abstracts, letters to journal editors and non-English language studies may have led to publication bias, an inflation of accuracy estimates due to preferential acceptance of papers reporting favourable results, and the potential for publication bias in studies included in the present meta-analysis was observed. Second, we did not address issues such as cost-effectiveness, reliability, the incremental benefit of adding tumour marker assays to other tests, and the net effect of tumour marker assays on clinical care and patient outcomes. Also, because of lack of required data reported in the original publications, we could not analyse the effect of factors such as laboratory infrastructure, expertise with tumour marker assay technology, patient spectrum and setting on the accuracy of the tumour marker measurements.

The accuracy of tumour marker determinations for MPE seems to be similar to that of conventional tests such as cytological examination with a high specificity and low sensitivity. This similarity might make tumour markers less useful in practice because they do not have test properties that complement the properties of conventional tests. Based on the findings in our meta-analysis, we cannot recommend using any one tumour marker alone for the diagnosis of MPE. However, it should be pointed out that, to date, there are insufficient related studies to evaluate the diagnostic accuracy of the combination of two or more tumour markers in MPE.

In conclusion, current evidence suggests that CA 15-3, CA 19-9 and CYFRA 21-1 are highly specific but insufficiently sensitive

to diagnose MPE, and the combination of two or more tumour markers seems to be more sensitive. Based on our data, we think that every patient with unexplained pleural effusion should undergo thoracocentesis with measurement of tumour markers. Patients with negative cytological examinations and positive tumour marker levels should undergo further invasive procedures, and management decisions should depend on positive cytological or biopsy results of the pleura.

REFERENCES

1. **American Thoracic Society**. Management of malignant pleural effusions. *Am J Respir Crit Care Med* 2000;**162**:1987–2001.
2. **Bennett R,** Maskell N. Management of malignant pleural effusions. *Curr Opin Pulm Med* 2005;**11**:296–300.
3. **Marel M,** Stastny B, Melinova L, et al. Diagnosis of pleural effusions. Experience with clinical studies, 1986 to 1990. *Chest* 1995;**107**:1598–603.
4. **Prakash UB,** Reiman HM. Comparison of needle biopsy with cytologic analysis for the evaluation of pleural effusion: analysis of 414 cases. *Mayo Clin Proc* 1985;**60**:158–64.
5. **Nance KV,** Shermer RW, Askin FB. Diagnostic efficacy of pleural biopsy as compared with that of pleural fluid examination. *Mod Pathol* 1991;**4**:320–4.
6. **Light RW.** Tumor markers in undiagnosed pleural effusions. *Chest* 2004;**126**:1721–2.
7. **Meyer T,** Rustin GJ. Role of tumor markers in monitoring epithelial ovarian cancer. *Br J Cancer* 2000;**82**:1535–8.
8. **Huang Y,** Zimmerman RL, Bibbo M. Diagnostic value of CA 15-3 antibody in detecting metastatic adenocarcinoma. *Anal Quant Cytol Histol* 2004;**26**:259–62.
9. **Molina R,** Santabarbara P, Filella X, et al. Relationship of CA 125 and CA 19.9 with lung carcinoma histological subtype: preliminary study. *Int J Biol Markers* 1989;**4**:215–20.
10. **Barak V,** Goike H, Panaretakis KW, et al. Clinical utility of cytokeratins as tumor markers. *Clin Biochem* 2004;**37**:529–40.
11. **Bossuyt PM,** Reitsma JB, Bruns DE, et al. Standards for Reporting of Diagnostic Accuracy. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 2003;**326**:41–4.
12. **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.
13. **Deville 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. **Jiang J,** Shi HZ, Liang QL, et al. Diagnostic value of interferon-γ in tuberculous pleurisy: a metaanalysis. Chest 2007;**131**:1133–41.
16. **Irwig L,** Macaskill P, Glasziou P, et al. Meta-analytic methods for diagnostic test accuracy. J Clin Epidemiol 1995;**48**:119–30.
17. **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.
18. **Suzuki S,** Moro-oka T, Choudhry NK. The conditional relative odds ratio provided less biased results for comparing diagnostic test accuracy in meta-analyses. J Clin Epidemiol 2004;**57**:461–9.
19. **Westwood ME,** Whiting PF, Kleijnen J. How does study quality affect the results of a diagnostic meta-analysis? BMC Med Res Methodol 2005;**8**:20.
20. **Egger M,** Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;**315**:629–34.
21. **Porcel JM,** Light RW. Diagnostic approach to pleural effusion in adults. Am Fam Physician 2006;**73**:1211–20.
22. **Glas AS,** Lijmer JG, Prins MH, et al. The diagnostic odds ratio: a single indicator of test performance. J Clin Epidemiol 2003;**56**:1129–35.
23. **Deeks JJ.** Systematic reviews of evaluations of diagnostic and screening tests. In: Egger M, Smith GD, Altman DG, eds. Systematic reviews in health care. Meta-analysis in context. London: BMJ Publishing Group, 2001:248–82.
24. **Jaeschke R,** Guyatt G, Lijmer J. Diagnostic tests. In: Guyatt G, Rennie D, eds. Users' guides to the medical literature. A manual for evidence-based clinical practice. Chicago: AMA Press, 2002:121–40.
25. **Booth SN,** Lakin G, Dykes PW, et al. Cancer-associated proteins in effusion fluids. J Clin Pathol 1977;**30**:537–40.
26. **Pettersson T,** Klockars M, Froseth B. Neuron-specific enolase in the diagnosis of small-cell lung cancer with pleural effusion: a negative report. Eur Respir J 1988;**1**:698–700.
27. **Petitti DB.** Approaches to heterogeneity in meta-analysis. Stat Med 2001;**20**:3625–33.

Pulmonary puzzle

A large chest wall tumour in an asymptomatic 15-year-old girl

CLINICAL PRESENTATION

A 15-year-old girl that had plain chest x ray as part of a routine medical examination was found to have a large left lower chest wall mass associated with rib destruction that was later shown to be painless and non-palpable (fig 1A). Chest CT revealed a 7×11 cm pleura based mass located in the left lower thorax and featured multiple rib destruction and downward depression of the left hemidiaphragm (fig 1B).

CT guided biopsy revealed a lesion with clusters of melanin–laden tumour cells. These tumour cells were positive for HMB-45 and S-100 protein immunohistochemically. The cytoplasmic pigments could be bleached by superoxidative agent and were positive by Fontana Masson stain, consistent with melanin pigments in nature. The skin overlying the tumour was intact, and the tumour was impalpable. Thorough skin examination over her whole body by an experienced dermatologist did not reveal any suspicious skin lesion. No other abnormality was demonstrated on whole body image studies.

QUESTION

What is the likely diagnosis and how should this be confirmed?
See page 94 for answer
This case is submitted by:

K-C Chen,[1] W-C Lin,[2] Y-C Lee,[1,3] J-S Chen,[1] H-H Hsu[1]

[1] Department of Surgery, National Taiwan University Hospital and National Taiwan University College of Medicine, Taipei, Taiwan; [2] Department of Pathology, National Taiwan University Hospital and National Taiwan University College of Medicine, Taipei, Taiwan; [3] Department of Traumatology, National Taiwan University Hospital and National Taiwan University College of Medicine, Taipei, Taiwan

Correspondence to: Dr H-H Hsu, Department of Surgery, National Taiwan University Hospital, No 7, Chung-Shan South Road, Taipei, Taiwan; ntuhfred@yahoo.com.tw
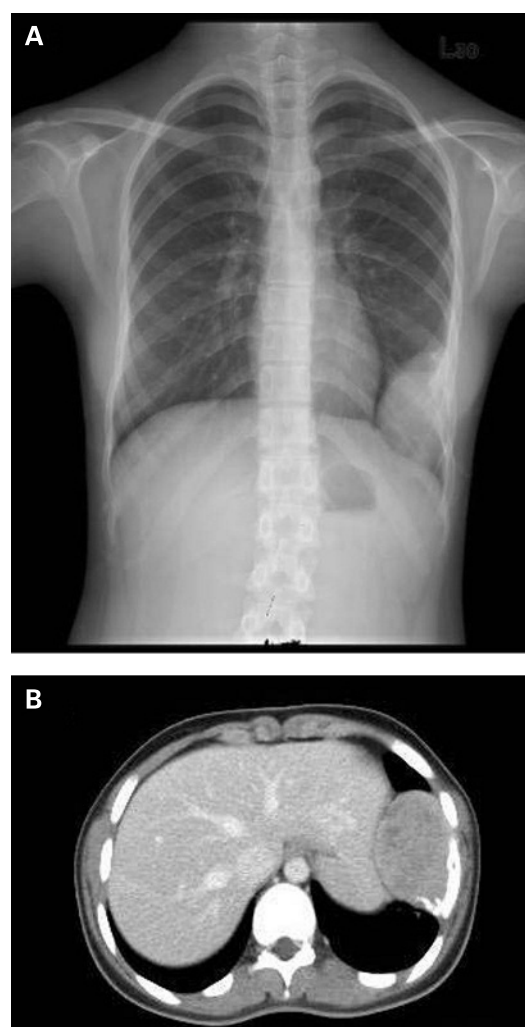
Competing interests: None.

Figure 1 (A) Chest plain film revealing a large, well defined left chest wall tumour with obvious rib destruction. (B) Thoracic CT revealing a large expansile tumour associated with osteolytic and sclerotic changes to a number of the patient's ribs.

Appendix statistical methods

For meta-analysis of diagnostic tests, a summary receiver operating characteristic (SROC) curve was constructed. We convert the true positive rate (TPR) and false positive rate (FPR) from each study to their logistic transforms. The method is based on the principle that there is a linear relationship between logit (TPR) and logit (FPR) where, using natural logs,

logit (TPR) = log (TPR / [1 − TPR]) and logit (FPR) = log (FPR / [1 − FPR]).

To estimate an SROC curve, we use the linear model:

$$D = \alpha + \beta S$$

Where

D = logit (TPR) − logit (FPR), S = logit (TPR) + logit (FPR), α = intercept, β = regression coefficient of S.

The model can be transformed back to the conventional axes of TPR against FPR, with SROC curves drawn only over the range of the data.

The model shown in the above equation ($D = \alpha + \beta S$) can be fit using conventional least squares methods (after adding 0.5 to each cell of the cross-classification of test and reference standard to deal with the possibility of zero cells) unweighted, i.e. giving equal weights to each study, or weighted by the inverse of the variance of D. Unweighted analysis has the obvious disadvantage of not paying more attention to larger studies, while weighted analysis may bias the estimate.

Appendix References

A1   Niwa Y, Shimokata K. Diagnostic significance of cancer antigen 125, pancreatic oncofetal antigen, and carcinoembryonic antigen in malignant and tuberculous pleural effusions. Jpn J Clin Oncol 1986; 16: 3–8.

A2   Lindgren J, Kuusela P, Hellstrom PE, et al. The ovarian cancer associated antigen CA 125 in patients with pleural effusions. Eur J Cancer Clin Oncol 1988; 24: 737–739.

A3   Shimokata K, Totani Y, Nakanishi K, et al. Diagnostic value of cancer antigen 15-3 (CA15-3) detected by monoclonal antibodies (115D8 and DF3) in exudative pleural effusions. Eur Respir J 1988; 1: 341–344.

A4   Lotzniker M, Pavesi F, Scarabelli M, et al. Tumour associated antigens CA 15.3 and CA 125 in ovarian cancer. Int J Biol Markers 1991; 6: 115–121.

A5   Satoh H, Sumi M, Yagyu H, et al. Clinical evaluation of CYFRA 21-1 in malignant pleural fluids. Oncology 1995; 52: 211–214.

A6   Romero S, Fernandez C, Arriero JM, et al. CEA, CA 15-3 and CYFRA 21-1 in serum and pleural fluid of patients with pleural effusions. Eur Respir J 1996; 9: 17–23.

A7   Toumbis M, Rasidakis A, Passalidou E, et al. Evaluation of CYFRA 21-1 in malignant and benign pleural effusions. Anticancer Res 1996; 16: 2101–2104.

A8   Villena V, Lopez-Encuentra A, Echave-Sustaeta J, et al. Diagnostic value of CA 72-4, carcinoembryonic antigen, CA 15-3, and CA 19-9 assay in pleural fluid. A study of 207 patients. Cancer 1996; 78: 736–740.

A9   San Jose ME, Alvarez D, Valdes L, et al. Utility of tumour markers in the diagnosis of

neoplastic pleural effusion. Clin Chim Acta 1997; 265: 193–205.

A10　Cynowska B, Słomiński JM, Goszka LW, et al. Value of determination of lung cancer marker Cyfra 21-1 for differentiation of pleural fluid – preliminary report. Med Sci Monit 1998; 4: 934–937.

A11　Ferrer J, Villarino MA, Encabo G, et al. Diagnostic utility of CYFRA 21-1, carcinoembryonic antigen, CA 125, neuron specific enolase, and squamous cell antigen level determinations in the serum and pleural fluid of patients with pleural effusions. Cancer 1999; 86: 1488–1495.

A12　Lai RS, Chen CC, Lee PC, et al. Evaluation of cytokeratin 19 fragment (CYFRA 21-1) as a tumor marker in malignant pleural effusion. Jpn J Clin Oncol 1999; 29: 421-424.

A13　Lee YC, Knox BS, Garrett JE. Use of cytokeratin fragments 19.1 and 19.21 (Cyfra 21-1) in the differentiation of malignant and benign pleural effusions. Aust N Z J Med 1999; 29: 765–769.

A14　Miedouge M, Rouzaud P, Salama G, et al. Evaluation of seven tumour markers in pleural fluid for the diagnosis of malignant effusions. Br J Cancer 1999; 81: 1059–1065.

A15　Kuralay F, Tokgöz Z, Cömlekci A. Diagnostic usefulness of tumour marker levels in pleural effusions of malignant and benign origin. Clinica Chimica Acta 2000; 300: 43–55.

A16　Alatas F, Alatas O, Metintas M, et al. Diagnostic value of CEA, CA 15-3, CA 19-9, CYFRA 21-1, NSE and TSA assay in pleural effusions. Lung Cancer 2001; 31: 9–16.

A17　Dejsomritrutai W, Senawong S, Promkiamon B. Diagnostic utility of CYFRA 21-1 in malignant pleural effusion. Respirology 2001; 6: 213–216.

A18 Paganuzzi M, Onetto M, Marroni P, et al. Diagnostic value of CYFRA 21-1 tumor marker and CEA in pleural effusion due to mesothelioma. Chest 2001; 119: 1138–1142.

A19 Lyubimova NV, Yag'ya TN, Chuchalin AG, et al. Diagnostic value of tumor markers Cyfra 21-1 and neuron-specific enolase in analysis of pleural fluid. Bull Exp Biol Med 2002; 133: 478–480.

A20 Sthaneshwar P, Yap SF, Jayaram G. The diagnostic usefulness of tumour markers CEA and CA-125 in pleural effusion. Malays J Pathol 2002; 24: 53–58.

A21 Hung TL, Chen FF, Liu JM, et al. Clinical evaluation of HER-2/neu protein in malignant pleural effusion-associated lung adenocarcinoma and as a tumor marker in pleural effusion diagnosis. Clin Cancer Res 2003; 9: 2605–2612.

A22 Jurman D, Lukinac L, Nöthig-Hus D, et al. Tumor marker CYFRA-21-1 in serum and pleural effusions of patients with lung cancer. Acta Clin Croat 2003; 42: 23–28.

A23 Villena V, Lopez-Encuentra A, Echave-Sustaeta J, et al. Diagnostic value of CA 549 in pleural fluid. Comparison with CEA, CA 15.3 and CA 72.4. Lung Cancer 2003; 40: 289–294.

A24 Porcel JM, Vives M, Esquerda A, et al. Use of a panel of tumor markers (carcinoembryonic antigen, cancer antigen 125, carbohydrate antigen 15-3, and cytokeratin 19 fragments) in pleural fluid for the differential diagnosis of benign and malignant effusions. Chest 2004; 126: 1757–1763.

A25 Tozkoparan E, Özkan M, Çiftçi F, et al. The role of serum and pleural fluid levels of Cyfra-21-1, carcinoembryonic antigen and neuron specific enolase in clinical evaluation of pleural effusions for malignancy. Archives of Lung 2004; 5: 194–199.

A26 Ustun H, Borazan A, Bilgicli N, et al. Diagnostic value of tumoural markers in pleural effusions. Int J Clin Pract 2004; 58: 22–25.

A27 Ghayumi SM, Mehrabi S, Doroudchi M, et al. Diagnostic value of tumor markers for differentiating malignant and benign pleural effusions of Iranian patients. Pathol Oncol Res 2005; 11: 236–241.

A28 Lee JH, Chang JH. Diagnostic utility of serum and pleural fluid carcinoembryonic antigen, neuron-specific enolase, and cytokeratin 19 fragments in patients with effusions from primary lung cancer. Chest 2005; 128: 2298–2303.

A29 Shitrit D, Zingerman B, Shitrit AB, et al. Diagnostic value of CYFRA 21-1, CEA, CA 19-9, CA 15-3, and CA 125 assays in pleural effusions: analysis of 116 cases and review of the literature. Oncologist 2005; 10: 501–507.

A30 Marukawa M, Hiyama J, Shiota Y, et al. The usefulness of CYFRA 21-1 in diagnosing and monitoring malignant pleural mesothelioma. Acta Med Okayama 1998; 52: 119–123.

A31 Terracciano D, Di Carlo A, Papa P, et al. New approaches in the diagnostic procedure of malignant pleural effusions. Oncol Rep 2004; 12: 79–83.

A32 Neskovic-Konstantinovic Z, Tasic S, Vuletic L, et al. CEA and CA 15-3 in pleural effusion of advanced breast cancer patients: clinical relevance and diagnostic value. Int J Biol Markers 1993; 8: 94–102.

A33 Cascinu S, Del Ferro E, Barbanti I, et al. Tumor markers in the diagnosis of malignant serous effusions. Am J Clin Oncol 1997; 20: 247–250.

A34 Mumbarkar PP, Raste AS, Ghadge MS. Significance of tumor markers in lung cancer.

Indian J Clin Biochem 2006; 21: 173–176.

A35 Pinto MM, Bernstein LH, Brogan DA, et al. Immunoradiometric assay of CA 125 in effusions. Comparison with carcinoembryonic antigen. Cancer 1987; 59: 218–222.

A36 Mezger J, Permanetter W, Gerbes AL, et al. Tumor associated antigens in diagnosis of serous effusions. J Clin Pathol 1988; 41: 633–643.

A37 Mezger J, Wilmanns W, Lamerz R. Elevated serum CA 125 levels in patients with benign ascitic or pleural effusion. Tumor Biol 1988; 9: 47–52.

A38 Ferroni P, Szpak C, Greiner JW, et al. CA 72-4 radioimmunoassay in the diagnosis of malignant effusions. Comparison of various tumor markers. Int J Cancer 1990; 46: 445–451.

A39 Kandylis K, Vassilomanolakis M, Baziotis N, et al. Diagnostic significance of the tumor markers CEA, CA 15-3 and CA 125 in malignant effusions in breast cancer. Ann Oncol 1990; 1: 435–438.

A40 Pinto MM, Bernstein LH, Rudolph RA, et al. Diagnostic efficiency of carcinoembryonic antigen and CA 125 in the cytological evaluation of effusions. Arch Pathol Lab Med 1992; 116: 626–231.

A41 Ammon A, Eiffert H, Reil S, et al. Tumor-associated antigens in effusions of malignant and benign origin. Clin Investig 1993; 71: 437–444.

A42 Trape J, Molina R, Sant F. Clinical evaluation of the simultaneous determination of tumor markers in fluid and serum and their ratio in the differential diagnosis of serous effusions. Tumour Biol 2004; 25: 276–281.

A43 Topalak O, Saygili U, Soyturk M, et al. Serum, pleural effusion, and ascites CA-125

levels in ovarian cancer and nonovarian benign and malignant diseases: a comparative study. Gynecol Oncol 2002; 85: 108–113.

A44 Niwa Y, Kishimoto H, Shimokata K. Carcinomatous and tuberculous pleural effusions. Comparison of tumor markers. Chest 1985; 87: 351–355.

A45 Salama G, Miedouge M, Rouzaud P, et al. Evaluation of pleural CYFRA 21-1 and carcinoembryonic antigen in the diagnosis of malignant pleural effusions. Br J Cancer 1998; 77: 472–476.

A46 Porcel JM, Salud A, Vives M, et al. Soluble oncoprotein 185^{HER-2} in pleural fluid has limited usefulness for the diagnostic evaluation of malignant effusions. Clin Biochem 2005; 38: 1031–1033.

Appendix Table 1 Characteristics of included studies

Study	Numbers Ma/non Ma	Reference standard	Cross-sectional design	Consecutive or random	Blinded design	Prospective	Reference No.*
Niwa, 1986	40/26	Histo/Cyto	Unknown	Unknown	Unknown	Unknown	A1
Lindgren, 1988	25/40	Histo/Cyto	Unknown	Unknown	Unknown	Unknown	A2
Shimokata, 1988	40/41	Histo/Cyto	Yes	Yes	Yes	Yes	A3
Lotzniker, 1991	22/17	Histo/Cyto	Unknown	Unknown	Unknown	Unknown	A4
Satoh, 1995	74/34	Histo/Cyto	No	Yes	Yes	No	A5
Romero, 1996	42/73	Histo/Cyto	Unknown	Yes	Yes	Yes	A6
Toumbis, 1996	115/99	Histo/Cyto	Unknown	Unknown	Unknown	Unknown	A7
Villena, 1996	65/142	Histo/Cyto	Unknown	Yes	Yes	Yes	A8
San Jose, 1997	88/183	Histo/Cyto	Unknown	Unknown	Unknown	Unknown	A9
Cynowska, 1998	22/13	Histo/Cyto	Unknown	Unknown	Unknown	Unknown	A10
Ferrer, 1999	43/79	Histo/Cyto	Unknown	Yes	Unknown	Yes	A11
Lai, 1999	72/54	Histo/Cyto	Unknown	Unknown	Unknown	No	A12
Lee, 1999	48/34	Histo/Cyto	Unknown	Yes	Unknown	Unknown	A13

Miedouge, 1999	215/121	Histo/Cyto	Yes	Yes	No	No	A14
Kuralay, 200	19/42	Histo/Cyto	Unknown	Unknown	Unknown	Unknown	A15
Alatas, 2001	44/30	Histo/Cyto	No	Yes	Yes	Yes	A16
Dejsomritrutai, 2001	27/35	Histo/Cyto	Yes	Yes	Yes	Yes	A17
Paganuzzi, 2001	72/34	Histo/Cyto	Yes	Yes	No	No	A18
Lyubimova, 2002	32/37	Histo/Cyto	No	Yes	No	Yes	A19
Sthaneshwar, 2002	25/53	Histo/Cyto	No	Yes	Yes	No	A20
Hung, 2003	50/95	Histo/Cyto	Yes	No	Yes	No	A21
Jurman, 2003	37/26	Histo/Cyto	Yes	Yes	Yes	No	A22
Villena, 2003	101/151	Histo/Cyto	Unknown	Yes	Yes	Yes	A23
Porcel, 2004	166/250	Histo/Cyto	Yes	Yes	Yes	Yes	A24
Tozkoparan, 2004	20/20	Histo/Cyto	Unknown	Unknown	Unknown	Unknown	A25
Ustun, 2004	41/40	Histo/Cyto	Unknown	No	Unknown	Yes	A26
Ghayumi, 2005	40/37	Histo/Cyto	Yes	Yes	Yes	Yes	A27
Lee, 2005	34/16	Histo/Cyto	No	Unknown	Unknown	Yes	A28
Shitrit, 2005	44/72	Histo/Cyto	Yes	Yes	Yes	Yes	A29

* Appendix References are available at http://thorax.bmj.com/supplemental

Ma = Malignant; Histo = Histology; Cyto = Cytology.

Appendix Table 2 Summary of CA 125 studies

Study	Numbers of patients	Assay method	Cut-off	TP	FP	FN	TN	Quality score		Reference No.*
								STARD	QUADAS	
Niwa, 1986	66	RIA	2100 U/ml	11	0	29	26	11	9	A1
Lindgren, 1988	65	RIA	250 U/ml	18	10	7	30	8	9	A2
Lotzniker, 1991	41	EIA	35 U/ml	22	18	0	1	9	9	A4
San Jose, 1997	271	EIA	518 U/ml	62	71	26	112	11	11	A9
Ferrer, 1999	122	RIA	1000 ng/ml	16	0	27	79	15	10	A11
Kuralay, 2000	61	EIA	352 U/ml	18	2	1	40	12	9	A15
Sthaneshwar, 2002	78	ECIA	1707 U/ml	9	3	16	50	14	11	A20
Porcel, 2004	416	EIA	2800 U/ml	28	0	138	250	16	12	A24
Ghayumi, 2005	77	EIA	1197 U/ml	19	11	21	26	14	12	A27
Shitrit, 2005	116	RIA	35 U/ml	43	4	1	68	15	11	A29

* Appendix References are available at http://thorax.bmj.com/supplemental

TP = true positive; FP = false positive; FN = false negative; TN = true negative; EIA = Enzyme immunoassay; ECIA =

Electrochemiluminescence immunoassay; RIA = Radioimmunoassay, STARD = standards for reporting diagnostic accuracy; QUADAS = quality assessment for studies of diagnostic accuracy.

Appendix Table 3 Summary of CA 15-3 studies

Study	Numbers of patients	Assay method	Cut-off	TP	FP	FN	TN	Quality score		Reference No.*
								STARD	QUADAS	
Shimokata, 1988	81	RIA	16 U/ml	15	0	25	41	10	12	A3
Lotzniker, 1991	39	EIA	25 U/ml	10	2	12	15	8	7	A4
Romero, 1996	115	EIA	25 U/ml	20	2	24	71	17	12	A6
Villena, 1996	207	RIA	42 U/ml	36	13	29	129	16	12	A8
Miedouge, 1999	336	RIA	36.2 U/ml	140	1	75	120	14	9	A14
Alatas, 2001	74	EIA	14 U/ml	35	2	9	28	15	13	A16
Villena, 2003	252	RIA	45 U/ml	44	1	57	150	16	12	A23
Porcel, 2004	416	ECIA	75 U/ml	50	0	116	250	16	12	A24
Ustun, 2004	81	EIA	30 U/ml	21	10	20	30	12	8	A26
Ghayumi, 2005	77	EIA	21.1 U/ml	28	6	12	31	14	12	A27
Shitrit, 2005	105	ECIA	30 U/ml	17	2	24	62	15	11	A29

* Appendix References are available at http://thorax.bmj.com/supplemental

TP = true positive; FP = false positive; FN = false negative; TN = true negative; EIA = Enzyme immunoassay; ECIA = Electrochemiluminescence immunoassay; RIA = Radioimmunoassay, STARD = standards for reporting diagnostic accuracy; QUADAS = quality assessment for studies of diagnostic accuracy.

Appendix Table 4 Summary of CA19-9 studies

Study	Numbers of patients	Assay method	Cut-off	TP	FP	FN	TN	Quality score		Reference No.*
								STARD	QUADAS	
Shimokata, 1988	91	RIA	10 U/ml	30	2	19	40	10	12	A3
Villena, 1996	207	RIA	580 U/ml	13	0	52	142	16	12	A8
Miedouge, 1999	336	RIA	219 U/ml	45	0	291	120	14	9	A14
Kuralay, 2000	61	EIA	54 U/ml	17	1	2	41	12	9	A15
Alatas, 2001	74	EIA	5 U/ml	16	5	28	25	15	13	A16
Ustun, 2004	81	EIA	33 U/ml	16	11	25	29	12	8	A26
Shitrit, 2005	116	EIA	30 U/ml	11	0	33	72	15	11	A29

* Appendix References are available at http://thorax.bmj.com/supplemental

TP = true positive; FP = false positive; FN = false negative; TN = true negative; EIA = Enzyme immunoassay; ECIA = Electrochemiluminescence immunoassay; RIA = Radioimmunoassay, STARD = standards for reporting diagnostic accuracy; QUADAS = quality assessment for studies of diagnostic accuracy.

Appendix Table 5 Summary of CYFRA 21-1 studies

Study	Numbers of patients	Assay method	Cut-off	TP	FP	FN	TN	Quality score		Reference No.*
								STARD	QUADAS	
Satoh, 1995	108	EIA	20.9 ng/ml	61	10	13	24	11	12	A5
Romero, 1996	115	EIA	50.0 ng/ml	16	13	26	60	17	12	A6
Toumbis, 1996	214	RIA	50.0 ng/ml	80	8	35	91	11	9	A7
Cynowska, 1998	35	RIA	3.3 ng/ml	20	12	2	1	10	9	A10
Ferrer, 1999	122	RIA	150 U/L	10	0	33	79	15	10	A11
Lai, 1999	126	RIA	50.0 ng/ml	41	3	31	51	11	9	A12
Lee, 1999	82	RIA	32.0 ng/ml	37	7	11	27	11	6	A13
Miedouge, 1999	336	EIA	163.0 ng/ml	92	1	123	120	14	9	A14
Alatas, 2001	74	EIA	8.0 ng/ml	40	3	4	27	15	13	A16
Dejsomritrutai, 2001	62	EIA	55.0 ng/ml	20	1	7	34	15	12	A17
Paganuzzi, 2001	106	EIA	41.9 ng/ml	56	7	16	27	14	9	A18
Lyubimova, 2002	69	EIA	75.0 ng/ml	14	1	18	36	12	9	A19

Hung, 2003	145	EIA	60.0 ng/ml	20	5	10	90	11	10	A21
Jurman, 2003	63	RIA	3.3 ng/ml	27	7	10	19	14	12	A22
Porcel, 2004	416	ECIA	175.0 ng/ml	37	0	144	250	16	12	A24
Tozkoparan, 2004	40	RIA	20.0 ng/ml	15	2	5	18	12	10	A25
Lee, 2005	50	RIA	45 ng/ml	21	3	13	13	12	9	A28
Shitrit, 2005	116	EIA	3.3 ng/ml	26	14	18	58	15	11	A29

* Appendix References are available at http://thorax.bmj.com/supplemental

TP = true positive; FP = false positive; FN = false negative; TN = true negative; EIA = Enzyme immunoassay; ECIA = Electrochemiluminescence immunoassay; RIA = Radioimmunoassay, STARD = standards for reporting diagnostic accuracy; QUADAS = quality assessment for studies of diagnostic accuracy.