DR. ILARIA GIROLAMI (Orcid ID : 0000-0003-3537-4929)
DR. ALBINO ECCHER (Orcid ID : 0000-0002-9992-5550)
DR. PASQUALE PISAPIA (Orcid ID : 0000-0002-6429-0620)
PROFESSOR GIANCARLO TRONCONE (Orcid ID : 0000-0003-1630-5805)
DR. LIRON PANTANOWITZ (Orcid ID : 0000-0001-8182-5503)

Article type : Review Article

Evidence-Based Diagnostic Performance of Novel Biomarkers for the Diagnosis of Malignant Mesothelioma in Effusion Cytology

Running title: Mesothelioma biomarkers in cytology

Ilaria Girolami MD[1], Ersilia Lucenteforte ScD[2], Albino Eccher MD[3], Stefano Marletta MD[4], Matteo Brunelli MD[4], Paolo Graziano MD[5], Pasquale Pisapia MD[6], Umberto Malapelle PhD[6], Giancarlo Troncone MD[6], Aldo Scarpa MD[3,4], Tao Huang MD[7], Liron Pantanowitz MD[7]

[1] Division of Pathology, Central Hospital Bolzano, Bolzano, Italy
[2] Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy
[3] Department of Pathology and Diagnostics, University and Hospital Trust of Verona
[4] Section of Pathology, Department of Diagnostics and Public Health, University and Hospital Trust of Verona, Verona, Italy
[5] Pathology Unit, Foundation IRCCS Casa Sollievo Della Sofferenza, San Giovanni Rotondo, Italy
[6] Department of Public Health, University of Naples Federico II, Naples, Italy
[7] Department of Pathology, University of Michigan, Ann Arbor, MI, USA

Address correspondence and reprint requests to: Albino Eccher, Department of Diagnostics and Public Health, University and Hospital Trust of Verona, P.le Stefani n. 1; 37126, Verona, Italy. Phone: +390458122161, Fax: +390458122011, e-mail: albino.eccher@aovr.veneto.it

Funding this research did not receive any specific grant from a funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest the authors declare that they do not have any conflict of interest.

Points for Table of Contents

- The diagnostic biomarkers BAP1 and MTAP by immunohistochemistry and *CDKN2A* (p16) by FISH, are reliable markers with superb specificity (close to 1.00) in distinguishing malignant from reactive mesothelial cells, while none of them is sufficient when utilized alone due to only moderate sensitivity (0.40-0.60).
- Diagnostic power is markedly improved when two of the contemporary biomarkers are used, such as a combination of BAP1 loss by immunohistochemistry and *CDKN2A* (p16) homozygous deletion by FISH. A combined analysis of BAP1 and MTAP loss by immunohistochemistry is also promising.

ABSTRACT

Background: Cytology effusions are often the only available material for diagnosing malignant pleural mesothelioma (MPM). However, the cytomorphological features alone are not always diagnostic, and cytology samples preclude assessment for pleural tissue invasion. Accordingly, immunohistochemical, soluble and molecular biomarkers have been developed. Aim of study was to provide quantitative evidence regarding the diagnostic performance for novel biomarkers.

Methods: A systematic literature review was carried out involving articles dealing with loss of BAP1, MTAP and 5-hmC, GLUT1, IMP3, EZH2 staining, *CDKN2A* homozygous deletion (HD) testing, soluble mesothelin and miRNAs quantification in cytological samples for the diagnosis of MPM versus reactive atypical mesothelial cells. Sensitivity and specificity were extracted and a meta-analysis performed. Quality of studies was assessed with QUADAS-2, and quality of evidence evaluated with GRADE approach.

Results: 71 studies were included. BAP1 loss showed sensitivity and specificity of 0.65 (CI 0.59-0.71) and 0.99 (CI 0.93-1.00), respectively. MTAP loss and p16 HD showed 100% specificity with sensitivities of 0.47 (CI 0.38-0.57) and 0.62 (CI 0.53-0.71), respectively. BAP1 loss and *CDKN2A* HD combined showed maximal specificity and sensitivity of 0.83 (CI 0.78-0.89). GLUT1 and IMP3 showed sensitivities of 0.82 (CI0.70-0.90) and 0.65 (CI 0.41-0.90) with comparable specificity, respectively. Mesothelin showed 0.73 (CI 0.68-0.77) sensitivity and 0.90 (CI 0.84-0.93) specificity.

Conclusion: Some of the recently emerged biomarkers are close to 1.00 specificity. Their moderate sensitivity however on their own can be significantly improved by performing two biomarkers, such as combining BAP1 with *CDKN2A* using FISH, or combining BAP1 and MTAP immunohistochemistry.

Keywords

Cytology, biomarker, diagnostic specificity and sensitivity, effusion, immunohistochemistry, mesothelioma, meta-analysis, pleura, systematic review

Counts

20 pages, 2 tables, 9 figures, 6 supporting files

BACKGROUND

Malignant mesothelioma arises from the serosal surfaces lining the pleural, peritoneal and pericardial cavities.[1] Exposure to asbestos fibers is regarded as the major etiological factor, but the role of genetic predisposition is increasingly recognized.[2] Malignant pleural mesothelioma (MPM) carries a poor prognosis, with an overall survival of less than 18 months.[1,3] Afflicted patients are usually elderly and not always fit enough to tolerate thoracoscopic surgery to obtain a pleural biopsy. Cytological examination of the pleural effusion on the other hand can be performed with minimal invasion with less morbidity. As tissue invasion is a major diagnostic criterion, mesothelioma guidelines recommend that cell block preparation should be performed whenever possible.[3] Prior published sensitivity for rendering a diagnosis of mesothelioma based on cytologic evaluation alone ranges from 0.30 to 0.75.[4] Although reliable immunohistochemical (IHC) markers have been well established to assist in the differential diagnosis between MPM and metastatic adenocarcinoma in pleural effusions,[5,6] a panel of biomarkers has not been established to reliably differentiate between MPM and reactive atypical mesothelial cells.

According to recent International Mesothelioma Interest Group guidelines,[7] the most valuable biomarkers to discriminate MPM from benign mesothelial proliferations include loss of IHC expression of BRCA1 associated protein 1 (BAP1) and the homozygous deletion (HD) of the cyclin dependent kinase inhibitor 2A (*CDKN2A)* gene (also known as *p16*), evaluated by fluorescent in situ hybridization (FISH). A meta-analysis to evaluate the diagnostic performance of BAP1 loss showed that the sensitivity by immunohistochemistry is 0.74 in epithelioid mesothelioma, 0.50 in biphasic mesothelioma, and 0.07 in sarcomatoid,[8,9] with confirmation in subsequent reports.[10,11] *CDKN2A* has the same unsatisfactory sensitivity (0.48-0.88), which is slightly higher (up to 0.80-1.00) for sarcomatoid mesothelioma,[10,12] and high specificity, as for recent reviews,[13,14] also in cytological material.[11]

Recently, methylthioadenosine (MTAP) IHC loss has emerged as another potentially useful biomarker. MTAP has shown high specificity for diagnosing MPM, with a sensitivity comparable to BAP1 and *CDKN2A* testing.[15] Other IHC biomarkers have also been tested such as insulin like-growth factor II messenger RNA-binding protein 3 (IMP3) and glucose transporter 1 (GLUT1). Despite some studies suggesting high specificity and variable sensitivity (0.30-0.75),[16,17] this is not truly established, with reporting of a large quota of benign lesions incorrectly staining.[18,19] Over-expression of Enhanced Zeste Homologue 2 (EZH2) appears to also play a role in mesothelioma development and progression and its IHC staining is nuclear, which allows it to be easily used in combination with other membranous or cytoplasmic markers. Early reports regarding EZH2 involved tissue specimens and showed a 1.00 specificity when 50% or high-staining/expression pattern is used as the cut off, with moderate sensitivity (0.45-0.66) when used alone, which increases in combination with BAP1/MTAP.[20,21] 5-hydroxymethylcitosine (5-hmC) is a modified nucleotide and its diagnostic use for pleural mesothelioma has only been reported in histological specimens.[22]

Soluble biomarkers include mesothelin, mesothelin-related peptides (SMRPs) and fibulin-3 detected by enzyme-linked immunosorbent assay (ELISA) or chemiluminescent enzyme immunoassay (CLEIA). They are relatively simple to detect and would permit a rapid diagnosis, but none of them has been proven to be highly sensitive to discriminate between benign and MPM.[18] Soluble mesothelin is the only Food and Drug Administration (FDA)-approved biomarker for MPM.[23] Elevated levels of mesothelin in pleural effusions have high specificity but low sensitivity,[23] and no unequivocal conclusions on its value has been reached.[24,25] Results on fibulin-3 are even more conflicting with this biomarker working better in plasma and with only one study dealing with pleural effusions.[26,27] MicroRNAs (miRNAs) are the newest potential diagnostic and prognostic markers. miRNAs are short RNA molecules involved in post-translational gene regulation and they have a well-established role in carcinogenesis, cancer progression, and metastasis.[28] miRNAs may be extracted from biological samples including liquid specimens such as plasma, saliva, and serous effusions.[29] However, the usefulness of miRNA profiling in effusions as a biomarker of MPM remains unknown.[30]

Aim of this study was to therefore conduct a systematic review to assess published evidence for the diagnostic accuracy of the aforementioned diagnostic biomarkers for the diagnosis of MPM in pleural effusion cytological material.

MATERIALS AND METHODS

A systematic review was conducted according to the guideline for Preferred Reporting Items for a Systematic review and Meta-Analysis of Diagnostic Test Accuracy (PRISMA-DTA).[31] The review was registered on the PROSPERO database (registration CRD42020198334).[32]

Search strategy

Electronic searches were carried out in the databases PubMed-MEDLINE, Embase and the Cochrane Library until 15th September, 2020 with separate searches for each biomarker. No study type filters were used nor language restriction applied. Where possible, filters to exclude animal and in vitro studies were applied. Search for grey literature was carried out in the opengrey.eu and oister.worldcat.org public resources. References listed in all included studies and previous reviews on biomarkers, even if not applicable to cytology, were hand-searched to retrieve potential additional studies. The complete search strategies are available in Supporting information, Table S1.

Article screening

Initial screening of articles by title/abstract was performed with aid of the online systematic review web-app QRCI.[33] Eligibility of published studies was determined independently by two reviewers with disagreement resolved through consensus. Following full-text screening, a list of excluded studies with reasons for exclusion was provided in a standardized PRISMA flow chart.

Eligibility criteria and data extraction

The criteria for including relevant studies were built following the model Population/Participants, Target condition, Index test, Reference test, Outcome, and Study design type. We considered studies dealing with patients with a pleural effusion and in which MPM was in the differential diagnosis with atypical mesothelial reactive cells. We considered only studies where the mesothelial origin of the atypical cellular component was already determined and an epithelial nature was excluded, thus excluding all studies dealing with the differential diagnosis of metastatic carcinoma. The Index test was represented by several biomarkers of interest. The Reference test was represented by a final diagnosis of MPM reached by means of definitive histology, or by means of follow-up long enough to demonstrate clinical evidence of MPM.

Data were extracted by a single reviewer with all outcomes data verified by a second reviewer, following the review question as for the protocol.[32]

The complete eligibility criteria for full-text inclusion and list of extracted data are reported in detail in Supporting information, Table S2.

Quality assessment of studies

The quality of included studies was assessed using the QUADAS-2 tool.[34] Risk of bias assessments was performed by one reviewer, with another reviewer providing subsequent verification. Briefly, QUADAS-2 tool comprises four domains for which basic yes/no questions guide the evaluation. The domains are: the patients' selection, which deals mainly with the avoidance of a case-control design (which is a source of bias) and the adoption of a

random/consecutive inclusion of patients; the index and references test domains, for which clear and exhaustive description with stating of the cut-off are required; and the flow and timing domain, with attention to the inclusion of all cases in the analysis and the timing of performing both index and reference test. Additionally, an evaluation of the applicability of the study to the review question is required. Risk of bias for all the domains and then overall judgement on single study is rated qualitatively low, high or unclear. Based on the QUADAS-2 guidance, we tailored the tool according to our review question. The adapted QUADAS-2 tool used is shown in Supporting information, Table S3.

Statistical analysis

Data from 2 x 2 tables for each biomarker (TP, FN, FP, TN) was used to summarize accuracy estimates. Graphical representation of studies was provided by plotting sensitivity and specificity estimates with their 95% confidence Intervals (CI) as a forest plot. As there was no explicit threshold reported for the biomarkers evaluated, a bivariate model was fitted using the METANDI function in STATA to calculate pooled estimates for sensitivity and specificity with their 95% CIs. For markers with very little variation in specificity, we fitted a univariate random model using the METAPROP function in STATA to calculate pooled estimates for sensitivity with its 95% CI. A subgroup analysis was performed according to geographical area of study (East versus Western countries), cell block versus smear, and risk of bias according to QUADAS2 (high versus low risk). All of the graphical depictions were performed with RevMan 5.3 (The Nordic Cochrane Centre, The Cochrane Collaboration, 2014) and STATA 15 (College Station, TX).

Quality of evidence assessment

For evaluation of the certainty of evidence of the pooled estimates of the diagnostic performance, the GRADE (Grading of Recommendations Assessment, Development and Evaluation) approach was used.[35] Briefly, GRADE is a rating system which applies to the outcome explored in review taking into account factors derived both from singles studies and from the final estimate from meta-analysis. The rating describes into four grades (very low, low, moderate and high) the overall quality of the final evidence obtained. The factors are the risk of bias of included studies, which concerns mainly the study design, the inconsistency of the estimate derived from visual inspection of forest plots showing not overlapping CI among singles studies, the indirectness for applicability of final estimate, when cases not entirely matching the population of interest, e.g. presence of data from biopsy not eliminable were present, the imprecision of the pooled estimate, mainly due to limited pooled sample size, and the publication bias. Given that all studies were likely to be observational and retrospective, certainty of evidence was always downgraded by at least one point for risk of bias due to study design. According to GRADE, judgment remains subjective but

explanation of decisions is required. Indeed, we were conservative and decided to take any issue of imprecision and inconsistency always as serious, downgrading further the quality of evidence.

RESULTS

Search algorithm

After duplicate removal, 2929 studies underwent title and abstract screening. Of these, 222 studies were checked in full-text form. There were 71 studies included in the qualitative synthesis, and 65 provided data for quantitative meta-analysis. The flowchart of screening articles is portrayed in Figure 1.

Study characteristics

The included studies comprised 57 full articles and 14 abstracts. The studies incorporated a total of 5354 patients from Europe (n = 25, 35%), North America (n = 19, 27%), Far East Asia (n = 16, 23%), Oceania (n = 6, 8%) and Middle East (n = 5, 7%). Regarding study design, 49 (69%) were retrospective, 19 (27%) collected cases prospectively and 3 (4%) declared a case-control design. Information on asbestos exposure was available in 7 (10%) studies. Detailed histological subtyping was available in 26 (37%) studies, with all comprising a majority of epithelioid MPM in their population; 3 studies dealt with epithelioid MPM only, 5 with both epithelioid and biphasic; none considered sarcomatoid type only. Cytological specimens were cell block only (n= 27, 38%), pleural effusion fluid for soluble markers (n = 26, 37%), cytological smears (n = 17, 24%), and a mixed population of smears and cell blocks (n = 1, 1%). Among soluble biomarkers, mesothelin/SMRP alone was investigated in 22 (31%) studies, fibulin-3 in 3 studies (4%) and both biomarkers together in one (1%). Among the IHC markers, BAP1 was investigated in 21 (30%) studies, MTAP in 8 (11%) studies, GLUT1 in 11 (16%) studies, and IMP3 in 5 (7%) studies. The newest IHC biomarkers EZH2 and 5-hmC were each investigated in one study. *CDKN2A* HD was investigated in 26 (37%) studies. Finally, miRNA signatures were investigated in 2 (3%) studies. The essential data of included studies are reported in Table 1, while the full list of included studies with all references is found in Supporting Table S4.

Quality assessment

The overall quality of studies was considered moderate. The parameter with the highest risk of bias was Patient selection in 14 (20%) studies, with some studies clearly declaring a case-control design, and almost all studies being observational retrospective. Critical points regarding the Index test domain concerned cut-offs in 12 (17%) studies with unclear or high risk of bias, while reference test was the issue with less risk of bias. Applicability concerns were generally limited,

with only few cases where presence of data from biopsy material or peritoneal specimen could not be excluded. The quality appraisal of studies is depicted in Figure 2 and Supporting Figure S1.

Diagnostic outcome – quantitative analysis

A bivariate model was used for the biomarkers BAP1 loss, p16 HD, GLUT1, IMP3, MTAP loss, mesothelin and the combination BAP1 loss and p16 HD. In the case of p16, IMP3 and the combination of BAP1 loss and p16 HD a univariate random model analysis was performed, as the specificity was always maximal at 1.00 or hampered by low variability around the maximal value. We also attempted a comparison among sensitivities of biomarkers taking the best-performing one, the combination of BAP1 loss and p16 HD, as reference, and we found that almost all markers significantly differed from the reference with lower sensitivity. Only GLUT1 did not show statistical significance for sensitivity, but its pooled specificity remained lower than that of the best-performing combination.

Concerning quality of evidence, at least one point of downgrading was always present due to retrospective and possibly case-control study design, with moderate as the highest grade, achieved by BAP1 loss, GLUT1, p16 HD and mesothelin/SMRP. Evidence judged of low (MTAP loss and combination of BAP1 loss and p16 HD) or very low (IMP3) quality were downgraded mainly for imprecision, due to limited pooled sample size, and/or inconsistency from inspection of forest plot.

The forest plots for single biomarker diagnostic performance with the relative ranges of sensitivity and specificity are shown in Figures 3 to Figure 9, while a summary of findings with highlighted the pooled estimates, comparisons and quality of evidence are revealed in Table 2.

Exploration of heterogeneity

Subgroup analysis was possible according to specimen type only for p16 HD. For the other biomarkers a minimum number of studies per subgroup was not reached.

The complete results of the subgroup analyses are reported in Supporting Information, Table S5. Concerning p16 HD, the pooled sensitivity was slightly higher with cytology smears than with cell blocks, but with no statistically significant difference with overall estimate. The pooled sensitivity and specificity of GLUT1 staining and BAP1 loss in cell block only cases showed substantial overlap with the overall estimates, and the same applied to the combination of BAP1 loss and p16 HD. Geographical area was not a moderator of heterogeneity. The extractable data did not allow for subgroup analysis according to histological subtyping nor to asbestos exposure, but we noticed that in all studies the absolute majority of MPM cases were of epithelioid subtype. Formal metrics such as Q and I^2 are not produced, since they are not appropriate for meta-analysis of diagnostic tests, given that they are univariate and sensitivity and specificity are correlated.[36]

Diagnostic outcome – qualitative synthesis

For the biomarkers fibulin-3, EZH2, 5-hmC and the miRNA signatures it was not possible to perform a meta-analysis. The sensitivity and specificity of fibulin-3 ranged from 0.79 to 0.88 and from 0.78 to 0.95, respectively in the two studies providing data.[37,38] The role of miRNAs were investigated in two studies,[39,40] but only one provided extractable data.[40] Other combinations that were investigated were EZH2 overexpression combined with BAP1 loss or p16 HD or MTAP loss or both,[41] BAP1 loss with MTAP loss,[42,43] and BAP1 loss with 5-mhC overexpression with or without MTAP loss.[44] The combination of BAP1 and MTAP loss with or without p16 HD and the combination of EZH2 overexpression with BAP1 or MTAP loss and p16 HD all showed 1.00 specificity, with sensitivity ranging from 0.72 to 0.91. 5-hmC loss alone or in combination with BAP1 loss showed a sensitivity and specificity ranging from 0.67 to 0.89 and from 0.89 to 1.00, respectively.

DISCUSSION

The cytological diagnosis of MPM typically cannot rely solely on the cytomorphological features given the significant overlapping cytologic features between reactive and malignant mesothelial proliferations. Previous systematic reviews attempting to summarize the diagnostic utility of biomarkers for MPM have failed to draw a strong conclusion. Our analysis indicates that BAP1 loss with IHC alone carries a high pooled specificity (0.99, CI 0.93-1.00), but still lower sensitivity (0.65, CI 0.59-0.71). This is in line with the only previous meta-analysis focused on cytological material.[8] *CDKN2A* HD showed comparable diagnostic performance with a pooled sensitivity of 0.62 (CI 0.53-0.71) and 1.00 specificity. These two biomarkers are diagnostically powerful in that a positive result from either of them is diagnostic of MPM. However, a negative result dose not exclude a diagnosis of MPM given their only moderate sensitivity when applied alone. Another important finding is that the combination of BAP1 loss and *CDKN2A* HD yields a pooled sensitivity of 0.83 (CI 0.78-0.89) with 1.00 specificity. The sensitivity is significantly increased with the combination (as shown by the CIs not overlapping) and this implies that, even though a negative result cannot exclude malignancy, a significantly greater quota of mesotheliomas can be detected with their combined use. Our subgroup analysis showed that the diagnostic performance of these two biomarkers was not significantly different from the overall estimate when considering cell block only cases, implying that even though cell block processing is always recommended, the diagnostic value is maintained irrespective of the specimen.

This study revealed interesting findings concerning the role of MTAP loss by IHC, which has shown high concordance with *CDKN2A*/p16 HD by FISH[45] and could potentially replace FISH analysis in low-resource settings. Unfortunately, the minimum number of studies to do a pooled analysis after combining BAP1 and MTAP loss was not reached, but in all three studies[41–43] that reported this combination there was a tendency toward improved sensitivity, ranging from 0.7 to 0.85. This

observation means that the addition of MTAP testing to BAP1 likely portends an improvement of sensitivity. It appears how the sensitivity of single marker remains suboptimal, or said otherwise, too different from specificity, and how it could be more advantageous to always combine the markers in order not to miss MPM cases without losing specificity and optimizing the diagnostic yield, in line with proposed diagnostic algorithms for tissue biopsy material.[10] Given that these markers are intended to aid and guide establishment of diagnosis and not as screening or triage tests, the search for higher sensitivity has to be balanced with preservation of higher specificity, so that a positive result confirms the diagnosis but without the risk of false-positive and overdiagnosis. Moreover, what is not always addressed in primary studies is the necessary presence of a positive internal control for markers such as BAP1 and MTAP which are deemed positive when showing loss of staining. All these markers maintain high specificity and this could be important also in case of a discordant result of two marker, as in case of a strongly suspicious case the repetition of the "discordant" marker could be performed or the addition of a third marker e.g. p16 after BAP1 and MTAP could counterbalance this situation.

This study also found a pooled sensitivity and specificity of 0.82 (CI 0.70-0.90) and 0.88 (CI 0.81-0.92) respectively for GLUT1 IHC staining. Our results are in agreement with a previous systematic research.[19] Overall, these values indicate an unfavorable diagnostic performance for GLUT1 suggesting that, even if an important quota of mesothelioma is correctly detected, there is likely to still be a proportion of negative cases which stain unreliably. Similar considerations apply to IMP3, which showed an unsatisfactory pooled sensitivity of 0.65 (CI 0.41-0.90). Therefore, IMP3 is not recommended to use for diagnosing MPM.

We also attempted to compare the performance of markers according to sensitivity, taking the best-performing one, the combination of BAP1 loss and p16 HD, as reference. Pooled sensitivity of other markers was significantly lower, implying a potential advantage of usage of this combination in detecting MPM. Only exception was GLUT1 which did not reach significance, but its pooled specificity was however lower than the reference. Moreover, these comparisons are statistically indirect, given that no primary studies compared a marker against another, so the comparative results should be regarded only as an indirect evaluation of their diagnostic performance.

Newer biomarkers such as EZH2 and 5-hmC showed a diagnostic profile similar to MTAP or BAP1 loss alone, and increase in sensitivity when used in combinations.[41] Due to the limited study, these two markers are not recommended for general use in diagnosing MPM in effusion cytology. Interestingly, these newer markers are deemed positive when overexpressed oppositely to MTAP and BAP1, without the need for an internal positive control, and this could be useful e.g. as third marker of a combination or in case of discordance in a pair, so we may expect that with future studies their use as adjunctive markers will increase. However, it is to be kept in mind that the marker itself or the combination does not make the diagnosis alone, but it always has to be evaluated together with morphology and clinical context.

No quantitative analysis was possible for miRNAs signatures with the two included studies.[39,40] Most of the studies about miRNAs signatures retrieved during the search process dealt with tissue specimens.[46] Finally, mesothelin showed pooled estimates for sensitivity and specificity of 0.73 (CI 0.68-0.77) and 0.90 (CI 0.84-0.93), respectively. Though these results should be interpreted with caution because there was evidence of threshold effect (correlation between logit sensitivity and logit specificity was -0.98), they are in keeping with results from previous systematic review.[25] This finding confirms the unsatisfactory profile of mesothelin/SMRPs in pleural effusion. Indeed, high specificity indicates that mesothelin could be helpful to confirm MPM, but a result below the cut-off for positivity does not exclude malignancy. Moreover, it must be kept in mind that elevate levels of mesothelin/SMRP are present also in other malignancies. It could be of interest to evaluate if the addition of mesothelin/SMRP measurement to a combination of two IHC markers or to the combination of BAP1 loss and p16 HD could further improve diagnostic yield, given that when pleural fluid is collected it can be easy and advantageous to perform different investigations on the same material. Unfortunately, no primary studies addressed this issue and this marker, even if performed together with others, has to be evaluated as a standalone test.

This study has both strengths and some limitations. The strengths reside in the methodology of performing a formal systematic review and the evaluation of the quality of evidence. The limitations of this study are related mainly to the primary studies with limited information included. A quota of risk of bias cannot be eliminated when studies declare a case-control design or are unclear when dealing with selection of cases or used different cut-offs for positivity or did not state clearly the distinction of target mesothelial cells for evaluation of staining or its loss. We chose to be conservative in judging quality of evidence, with all findings downgraded at least for one point for this reason. Moreover, for some markers quality of evidence was downgraded for imprecision due to limited pooled sample size and/or inconsistency from inspection of forest plot. We had to balance pros and cons keeping all the studies with a minimum of available quantitative information to maximize number of cases, balancing with the consideration that quality of evidence would be critically evaluated in light of this choice. This leads also to future directions for research emerged from this systematic review: given that the case-control design and selection of cases are the main sources of bias, prospective studies or even retrospective with uncontrolled selection could allow to draw evidence of higher quality.

In conclusion, our systematic review highlights how immunohistochemistry showing BAP1 loss, MTAP loss, and/or p16 HD by FISH have high specificity but suboptimal sensitivity when used alone. A combination of BAP1 loss and p16 HD yields a significant increase in MPM detection capability, making these dual biomarkers suitable to render a definitive diagnosis. Historical biomarkers such as GLUT1 and IMP3 as well as mesothelin/SMRPs when utilized alone have an unsatisfactory diagnostic performance and should not be relied upon for diagnosing MPM in pleural effusions.

REFERENCES

1. Robinson BW, Musk AW, Lake RA. Malignant mesothelioma. *Lancet*. 2005;366(9483):397-408. doi:10.1016/S0140-6736(05)67025-0
2. Cadby G, Mukherjee S, Musk AW (Bill), et al. A genome-wide association study for malignant mesothelioma risk. *Lung Cancer*. 2013;82(1):1-8. doi:10.1016/j.lungcan.2013.04.018
3. Scherpereel A, Opitz I, Berghmans T, et al. ERS/ESTS/EACTS/ESTRO guidelines for the management of malignant pleural mesothelioma. *Eur Respir J*. 2020;55(6). doi:10.1183/13993003.00953-2019
4. Henderson DW, Reid G, Kao SC, van Zandwijk N, Klebe S. Challenges and controversies in the diagnosis of malignant mesothelioma: Part 2. Malignant mesothelioma subtypes, pleural synovial sarcoma, molecular and prognostic aspects of mesothelioma, BAP1, aquaporin-1 and microRNA. *J Clin Pathol*. 2013;66(10):854-861. doi:10.1136/jclinpath-2013-201609
5. Hjerpe A, Ascoli V, Bedrossian C, et al. Guidelines for cytopathologic diagnosis of epithelioid and mixed type malignant mesothelioma. Complementary statement from the International Mesothelioma Interest Group, also endorsed by the International Academy of Cytology and the Papanicolaou Society o. *Cytojournal*. 2015;12:26. doi:10.4103/1742-6413.170726
6. Nottegar A, Tabbò F, Luchini C, et al. Pulmonary Adenocarcinoma With Enteric Differentiation. *Appl Immunohistochem Mol Morphol*. Published online October 2016:1. doi:10.1097/PAI.0000000000000440
7. Husain, AN, Colby, T V., Ordóñez, NG, et al. Guidelines for Pathologic Diagnosis of Malignant Mesothelioma 2017 Update of the Consensus Statement From the International Mesothelioma Interest Group. *Arch Pathol Lab Med*. 2018;142(1):89-108. doi:10.5858/arpa.2017-0124-RA
8. Wang L-M, Shi Z-W, Wang J-L, et al. Diagnostic accuracy of BRCA1-associated protein 1 in malignant mesothelioma: a meta-analysis. *Oncotarget*. 2017;8(40):68863-68872. doi:10.18632/oncotarget.20317
9. Mlika M, Zorgati M, BenKhelil M, Mezni F El. About the diagnostic value of BAP-1 antibody in malignant pleural mesothelioma: a meta-analysis. *J Immunoass Immunochem*. 2019;40(3):269-282. doi:10.1080/15321819.2019.1574814
10. Chapel DB, Schulte JJ, Husain AN, Krausz T. Application of immunohistochemistry in diagnosis and management of malignant mesothelioma. *Transl Lung Cancer Res*. 2020;9(S1):S3-S27. doi:10.21037/tlcr.2019.11.29
11. Churg A, Naso JR. The Separation of Benign and Malignant Mesothelial Proliferations: New Markers and How to Use Them. *Am J Surg Pathol*. Published online August 2020.

doi:10.1097/PAS.0000000000001565

12. Wu D, Hiroshima K, Matsumoto S, et al. Diagnostic Usefulness of p16/CDKN2A FISH in Distinguishing Between Sarcomatoid Mesothelioma and Fibrous Pleuritis. *Am J Clin Pathol*. 2013;139(1):39-46. doi:10.1309/AJCPT94JVWIHBKRD
13. Wan C, Shen Y-C, Liu M-Q, et al. Diagnostic Value of Fluorescence in Situ Hybridization Assay in Malignant Mesothelioma: A Meta-analysis. *Asian Pacific J Cancer Prev*. 2012;13(9):4745-4749. doi:10.7314/APJCP.2012.13.9.4745
14. Eccher A, Girolami I, Lucenteforte E, Troncone G, Scarpa A, Pantanowitz L. Diagnostic mesothelioma biomarkers in effusion cytology. *Cancer Cytopathol*. Published online January 19, 2021:cncy.22398. doi:10.1002/cncy.22398
15. Churg A, Nabeshima K, Ali G, Bruno R, Fernandez-Cuesta L, Galateau-Salle F. Highlights of the 14th international mesothelioma interest group meeting: Pathologic separation of benign from malignant mesothelial proliferations and histologic/molecular analysis of malignant mesothelioma subtypes. *Lung Cancer*. 2018;124:95-101. doi:10.1016/j.lungcan.2018.07.041
16. Minato H, Kurose N, Fukushima M, et al. Comparative immunohistochemical analysis of IMP3, GLUT1, EMA, CD146, and desmin for distinguishing malignant mesothelioma from reactive mesothelial cells. *Am J Clin Pathol*. 2014;141(1):85-93. doi:10.1309/AJCP5KNL7QTELLYI
17. Chang S, Oh M-H, Ji S-Y, et al. Practical utility of insulin-like growth factor II mRNA-binding protein 3, glucose transporter 1, and epithelial membrane antigen for distinguishing malignant mesotheliomas from benign mesothelial proliferations. *Pathol Int*. 2014;64(12):607-612. doi:10.1111/pin.12216
18. Churg A, Sheffield BS, Galateau-Salle F. New Markers for Separating Benign From Malignant Mesothelial Proliferations: Are We There Yet? *Arch Pathol Lab Med*. 2016;140(4):318-321. doi:10.5858/arpa.2015-0240-SA
19. Zhong S-C, Ao X-J, Yu S-H. Diagnostic value of GLUT-1 in distinguishing malignant mesothelioma from reactive mesothelial cells: a meta-analysis. *Biomarkers Biochem Indic Expo response, susceptibility to Chem*. 2020;25(2):157-163. doi:10.1080/1354750X.2020.1714735
20. Yoshimura M, Kinoshita Y, Hamasaki M, et al. Highly expressed EZH2 in combination with BAP1 and MTAP loss, as detected by immunohistochemistry, is useful for differentiating malignant pleural mesothelioma from reactive mesothelial hyperplasia. *Lung Cancer*. 2019;130:187-193. doi:10.1016/j.lungcan.2019.02.004
21. Shinozaki-Ushiku A, Ushiku T, Morita S, Anraku M, Nakajima J, Fukayama M. Diagnostic utility of BAP1 and EZH2 expression in malignant mesothelioma. *Histopathology*. 2017;70(5):722-733. doi:10.1111/his.13123

22. Chapel DB, Husain AN, Krausz T. Immunohistochemical evaluation of nuclear 5-hydroxymethylcytosine (5-hmC) accurately distinguishes malignant pleural mesothelioma from benign mesothelial proliferations. *Mod Pathol an Off J United States Can Acad Pathol Inc*. 2019;32(3):376-386. doi:10.1038/s41379-018-0159-7
23. Creaney J, Robinson BWS. Malignant Mesothelioma Biomarkers: From Discovery to Use in Clinical Practice for Diagnosis, Monitoring, Screening, and Treatment. *Chest*. 2017;152(1):143-149. doi:10.1016/j.chest.2016.12.004
24. Cui A, Jin X-G, Zhai K, Tong Z-H, Shi H-Z. Diagnostic values of soluble mesothelin-related peptides for malignant pleural mesothelioma: updated meta-analysis. *BMJ Open*. 2014;4(2):e004145. doi:10.1136/bmjopen-2013-004145
25. Gao R, Wang F, Wang Z, et al. Diagnostic value of soluble mesothelin-related peptides in pleural effusion for malignant pleural mesothelioma. *Medicine (Baltimore)*. 2019;98(14):e14979. doi:10.1097/MD.0000000000014979
26. Pei D, Li Y, Liu X, et al. Diagnostic and prognostic utilities of humoral fibulin-3 in malignant pleural mesothelioma: Evidence from a meta-analysis. *Oncotarget*. 2017;8(8):13030-13038. doi:10.18632/oncotarget.14712
27. Ren R, Yin P, Zhang Y, et al. Diagnostic value of fibulin-3 for malignant pleural mesothelioma: A systematic review and meta-analysis. *Oncotarget*. 2016;7(51):84851-84859. doi:10.18632/oncotarget.12707
28. Berindan-Neagoe I, Monroig P del C, Pasculli B, Calin GA. MicroRNAome genome: A treasure for cancer diagnosis and therapy. *CA Cancer J Clin*. 2014;64(5):311-336. doi:10.3322/caac.21244
29. Rossi ED, Bizzarro T, Martini M, et al. The evaluation of miRNAs on thyroid FNAC: the promising role of miR-375 in follicular neoplasms. *Endocrine*. 2016;54(3):723-732. doi:10.1007/s12020-016-0866-0
30. Nicolè L, Cappello F, Cappellesso R, VandenBussche CJ, Fassina A. MicroRNA profiling in serous cavity specimens: Diagnostic challenges and new opportunities. *Cancer Cytopathol*. 2019;127(8):493-500. doi:10.1002/cncy.22143
31. McInnes MDF, Moher D, Thombs BD, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies. *JAMA*. 2018;319(4):388. doi:10.1001/jama.2017.19163
32. Girolami I, Eccher A, Lucenteforte E, Pantanowitz L. PROSPERO International prospective register of systematic reviews. Diagnostic accuracy of immunoistochemical, soluble and molecular markers in the differential diagnosis of malignant mesothelioma from benign mesothelial cells in pleural effusion cytology. Published 2020. https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=198334
33. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for

systematic reviews. *Syst Rev*. 2016;5(1):210. doi:10.1186/s13643-016-0384-4

34. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536. doi:10.7326/0003-4819-155-8-201110180-00009
35. Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011;64(4):401-406. doi:10.1016/j.jclinepi.2010.07.015
36. McGrath TA, Alabousi M, Skidmore B, et al. Recommendations for reporting of systematic reviews and meta-analyses of diagnostic test accuracy: a systematic review. *Syst Rev*. 2017;6(1):194. doi:10.1186/s13643-017-0590-8
37. Agha MA, El-Habashy MM, El-Shazly RA. Role of fibulin-3 in the diagnosis of malignant mesothelioma. *Egypt J Chest Dis Tuberc*. 2014;63(1):99-105. doi:10.1016/j.ejcdt.2013.10.004
38. Pass HI, Levin SM, Harbut MR, et al. Fibulin-3 as a Blood and Effusion Biomarker for Pleural Mesothelioma. *N Engl J Med*. 2012;367(15):1417-1427. doi:10.1056/NEJMoa1115050
39. Birnie KA, Prêle CM, Musk AWB, et al. MicroRNA Signatures in Malignant Pleural Mesothelioma Effusions. *Dis Markers*. 2019;2019:8628612. doi:10.1155/2019/8628612
40. Cappellesso R, Nicolè L, Caroccia B, et al. Young investigator challenge: MicroRNA-21/MicroRNA-126 profiling as a novel tool for the diagnosis of malignant mesothelioma in pleural effusion cytology. *Cancer Cytopathol*. 2016;124(1):28-37. doi:10.1002/cncy.21646
41. Yoshimura M, Hamasaki M, Kinoshita Y, et al. Utility of highly expressed EZH2 in pleural effusion cytology for the diagnosis of mesothelioma. *Pathol Int*. Published online July 29, 2020:pin.12990. doi:10.1111/pin.12990
42. Kinoshita Y, Hida T, Hamasaki M, et al. A combination of MTAP and BAP1 immunohistochemistry in pleural effusion cytology for the diagnosis of mesothelioma. *Cancer Cytopathol*. 2018;126(1):54-63. doi:10.1002/cncy.21928
43. Hiroshima K, Wu D, Hamakawa S, et al. HEG1, BAP1, and MTAP are useful in cytologic diagnosis of malignant mesothelioma with effusion. *Diagn Cytopathol*. 2020;(May):dc.24475. doi:10.1002/dc.24475
44. Shahi M, Antic T, Fitzpatrick C, Husain A, Krausz T. A combination of BAP1, 5-HMC and MTAP immunohistochemical staining in malignant mesothelioma effusions. *Mod Pathol*. 2020;33(3):420-430. doi:10.1038/s41379-019-0354-1
45. Chapel DB, Schulte JJ, Berg K, et al. MTAP immunohistochemistry is an accurate and reproducible surrogate for CDKN2A fluorescence in situ hybridization in diagnosis of malignant pleural mesothelioma. *Mod Pathol an Off J United States Can Acad Pathol Inc*. 2020;33(2):245-254. doi:10.1038/s41379-019-0310-0
46. Micolucci L, Akhtar MM, Olivieri F, Rippo MR, Procopio AD. Diagnostic value of microRNAs

in asbestos exposure and malignant mesothelioma: systematic review and qualitative meta-analysis. *Oncotarget*. 2016;7(36):58606-58637. doi:10.18632/oncotarget.9686

Figure Legends

Figure 1. Search flow according to the PRISMA diagram.

Figure 2. Quality assessment of included studies according to QUADAS-2.

Figure 3. Forest plot of BAP loss.

Figure 4. Forest plot of GLUT1 staining.

Figure 5. Forest plot of MTAP loss.

Figure 6. Forest plot of IMP3 staining.

Figure 7. Forest plot of p16 homozygous deletion HD with FISH.

Figure 8. Forest plot of BAP loss combined with p16 homozygous deletion.

Figure 9. Forest plot of soluble mesothelin/SMRPs.

Table 1. Summary of included studies.[¶]

Author, year (Country)[¶]	Sample size	Specimen	MPM subtype	Diagnostic test	Cut-off for positivity
Agha, 2014 (Egypt)[1]	34	PE	10 E, 6 S, 9 B	Fibulin-3	127,5 ng/mL
Agrawal, 2019 (USA)*[2]	33	PE	NS	BAP1	NS
Aleman, 2009 (Spain)[3]	39	PE	NS	SMRP/mesothelin	6 nmol/L
Amany, 2013 (Egypt)[4]	40	PE	NS	SMRP/mesothelin	3,5 nmol/L
Andrici, 2015 (USA)[5]	168	CB	NS	BAP1	Retention of staining up to 5% of presumed target cells
Battolla, 2017 (Italy)[#6]	97	PE	22 E, 4 S, 3 B	Fibulin-3	0,183 ng/ml
				SMRP/mesothelin	0,3 nM/L
Battolla, 2012 (Italy)[7]	181	PE	35 E, 9 S, 4 B	SMRP/mesothelin	9,3 nmol/L
Berg, 2020 (USA)[#8]	39	CB	NS	BAP1	100% loss; positive control present
				MTAP	Less than 25% cells with staining
Birnie, 2019 (USA)[#9]	36	PE	17 E, 2 S, 1 B	miRNAs	Variable
Blanquart, 2012 (France)[10]	76	PE	49 E, 4 S, 4 B	SMRP/mesothelin	14,6 nM/L
Bradley, 2013 (UK)*[11]	18	PE	NS	SMRP/mesothelin	NS
Bruno, 2019 (Italy)[12]	27	CB	NS	BAP1	100% loss in atypical mesothelial

					cells in presence of positive control
				p16	More than 11% HD in atypical mesothelial cells
				BAP1 + p16	-
Canessa, 2013 (Italy)*[13]	86	PE	NS	SMRP/mesothelin	20 nM/L
Canessa, 2013 (Italy)[14]	104	PE	25 E, 9 S 0 B	SMRP/mesothelin	19,6 nM/L
Canessa, 2012 (Italy)[15]	181	PE	35 E, 9 S, 4 B	SMRP/mesothelin	9,30 nM/L
Cappellesso, 2016 (Italy)[16]	53	PE	NS	miRNAs	0,49 folds
Chen, 2020 (China)*#[17]	110	CB	NS	BAP1	NS
				p16	NS
Cigognetti, 2015 (Italy)[18]	70	CB	NS	BAP1	NS
Cozzi, 2017 (Italy)[19]	114	CB	NS	BAP1	NS; all-or-nothing staining with internal control present
Creaney, 2014 (Australia)[20]	829	PE, CB	59 E, 19 S, 23 B	SMRP/mesothelin	20 nM/L
Creaney, 2013 (Australia)[21]	98	PE	NS	SMRP/mesothelin	20 nM/L
Creaney, 2007 (Australia)[22]	136	PE		SMRP/mesothelin	20 nM/L
Dagli, 2011 (Turkey)*[23]	20	CB	NS	GLUT1	NS
				IMP3	NS
Davies, 2009 (UK)[24]	99	PE	11 E, 5 S, 4 B	SMRP/mesothelin	20 nM/L

Deng, 2013 (USA)*[25]	32	CB	NS	GLUT1	More than 5% of target cells
	32			IMP3	More than 5% of target cells
Ferro, 2013 (Italy)[26]	79	PE	NS	SMRP/mesothelin	12,7 nM/L
Flores-Staino, 2010 (Sweden)[27]	39	PE	NS	p16	12 nuclei with deletion
Fujimoto, 2010 (Japan)[28]	49	PE	15 E, 4 S, 2 B	SMRP/mesothelin	8 nM/L
Galateau-Salle, 2008 (France)*[29]	37	CB		p16	NS
Hanley, 2007 (USA)[30]	26	CB	NS	IMP3	NS
Hasteh, 2010 (USA)[31]	58	CB	NS	GLUT1	>20% mesothelial cells with membrane staining
Hamasaki, 2012 (Japan)*[32]	13	PE	NS	p16	More than 10% HD
Hamasaki, 2019 (Japan)*#[33]	74	CB	NS	Combination: p16, MTAP, BAP1 and NF2	NS
Hatem, 2018 (USA)[34]	30	CB	13 E, 2 S, 3 B	BAP1	More than 50% loss of staining
Hida, 2015 (Japan)[35]	45	PE	NS	p16	More than 10% HD
Hiroshima, 2016 (Japan)[36]	39	CB	15 E, 1 S, 6 B	p16	More than 15% HD
Hiroshima, 2020 (Japan)[37]	67	CB	24 E, 9 B	HEG1	score >2
				BAP1	NS
				MTAP	NS
				BAP1 + MTAP	-

Hooper, 2012 (UK)[38]	54	PE	23 E, 3 S, 2 B	SMRP/mesothelin	20 nM/L
Hwang, 2016 (Canada)[39]	16	PE	NS	BAP1	NS
				p16	More than 12% HD
				BAP1 + p16	
Ikeda, 2010 (Japan)[40]	50	PE	NS	IMP3	NS
Ikeda, 2011 (Japan)[41]	61	PE	11 E	IMP3	NS
				GLUT1	NS
				GLUT1	NS
Illei, 2003 (USA)[42]	32	PE	NS	p16	More than 15% HD
Javadi, 2020 (Sweden)[#43]	82	PE	NS	SMRP/mesothelin	NS
Kee, 2010 (New Zealand)[44]	34	CB	15 E, 1 S, 1 B	GLUT1	NS
Kinoshita, 2018 (Japan)[45]	66	CB	NS	BAP1	More than 50% loss
				MTAP	More than 50% loss
				p16	More than 10% HD
				BAP1 + MTAP	-
				BAP1 + p16	-
Kinoshita, 2020 (Japan)*[46]	42	CB	NS	p16	NS
				MTAP	NS
				BAP1	NS

Kirschner, 2015 (Australia)[#47]	90	PE	27 E, 1 S, 2 B	Fibulin-3	NS
Kuperman, 2013 (USA)[48]	88	CB	NS	GLUT1	Any positivity
Leong, 2015 (Australia)[49]	37	PE	NS	SMRP/mesothelin	20 nM/L
Matsumoto, 2013 (Japan)[50]	35	PE	NS	p16	More than 10% HD
Matsumoto, 2019 (Japan)[51]	88	PE	NS	p16	More than 10% HD
				BAP1	More than 50% loss
				BAP1 + p16	-
McCroskey, 2017 (USA)[52]	32	CB	19 E	BAP1	Retention of staining up to 5% of presumed target cells
Mutlu, 2012 (Turkey)[53]	40	CB	19 E, 1 B	GLUT1	More than 10%
Nabeshima, 2012 (Japan)[54]	20	PE	NS	p16	More than 10% HD
Önder, 2019 (Turkey)[55]	46	PE, CB	NS	BAP1	100% loss
				GLUT1	More than 1%
Onofre, 2008 (Germany)[56]	72	PE	NS	p16	More than 5 nuclei
Pass, 2008 (USA)[57]	72	PE	NS	SMRP/mesothelin	12,6 nM/L
Pass, 2012 (USA)[58]		PE		SMRP/mesothelin	378,33 ng/mL
Pinheiro, 2012 (Portugal)[59]	20	PE	NS	GLUT1	combined score >2
Raza, 2020 (USA)*[60]	33	PE	11 E, 2 B	BAP1	NS
				MTAP	NS

Savic, 2010 (Switzerland)[61]	80	PE	44 E, 1 S, 7 B	p16	More than 15% HD
Scherpereel, 2006 (France)[62]	64	PE	NS	SMRP/mesothelin	10,4 nM/L
Schürch, 2018 (Switzerland)[63]	148	CB	NS	BAP1	100% loss
				GLUT1	Any positivity
Shahi, 2020 (USA)*[64]	108	CB	62 E, 20 B	BAP1	NS
				MTAP	NS
				5-hmC	NS
				BAP1 + 5-hmC	-
				BAP1 + 5-hmC + MTAP	-
Shen, 2009 (USA)[65]	73	CB	28 E	GLUT1	Any positivity
					Any positivity
Stockhammer, 2020 (Germany)[66]	72	PE	35 E, 6 S, 7 B	SMRP/mesothelin	13,1 nM/L
Vigani, 2011 (Italy)*[67]	140	PE	NS	SMRP/mesothelin	10,8 nM/L
Walts, 2016 (USA)[68]	63	CB	25 E, 6 B	BAP1	More than 50% loss with internal control present
				p16	More than 15% HD
				BAP1 + p16	-
Yamada, 2011 (Japan)[69]	69	PE	37 E, 5 S, 3 B	SMRP/mesothelin	10 nmol/L

Yoshimura, 2020 (Japan)[70]	60	CB	NS	BAP1	More than 50% loss
				MTAP	More than 50% loss
				EZH2	More than 50% expression
				p16	10% HD
				BAP1 + EZH2	-
				MTAP + EZH2	-
				p16 + EZH2	-
				BAP1 + MTAP	-
				BAP1 + p16	-
				BAP1 + MTAP/p16 + EZH2	-
Zhu, 2020 (USA)*[71]	59	CB	NS	MTAP	NS

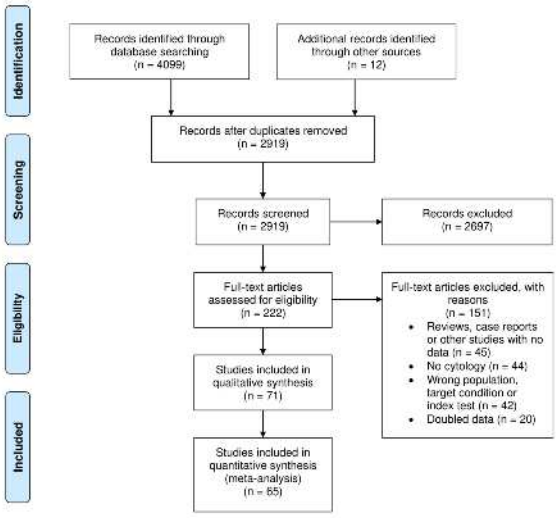
¶studies' numbers are as reported in Supplementary Table S4 for brevity reasons; *studies represented by abstracts only; #studies not providing data for quantitative analysis

Abbreviations: B, biphasic type; CB, cell block; E, epithelioid type; HD, homozygous deletion; NS, not stated; PE, pleural effusion (fluid or smear); S, sarcomatoid type

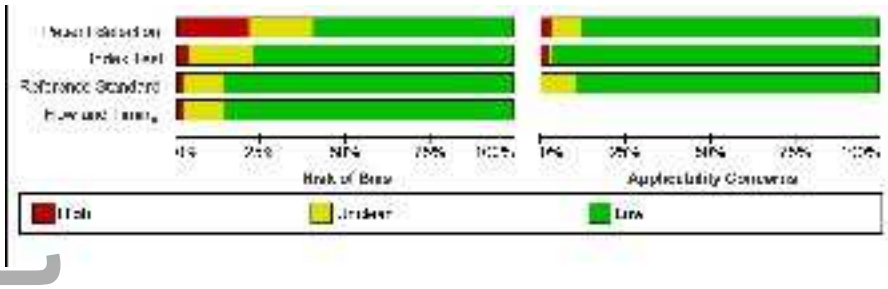
Table 2. Summary of sensitivity and specificity ranges and pooled estimates with comparisons and certainty of evidence.

Marker	Sensitivity range	Specificity range	Specificity – pooled estimate (95% CI)	Sensitivity – pooled estimate (95% CI)	Univariate comparison of sensitivities, p-value	Certainty of evidence - summary
BAP1 loss + p16 HD	0.63-0.92	1.00	1.00*	0.83 (0.78 to 0.89)*	(reference)	⊕⊕◯◯ - LOW
BAP1	0.41-0.86	0.96-1.00	0.99 (0.93-1.00)	0.65 (0.59 to 0.71)	0.00	⊕⊕⊕◯ - MODERATE
GLUT1	0.47-1.00	0.67-1.00	0.88 (0.81-0.92)	0.82 (0.80 to 0.90)	0.29	⊕⊕⊕◯ - MODERATE
MTAP	0.23-0.76	0.93-1.00	0.99 (0.88-1.00)	0.47 (0.38 to 0.57)	0.00	⊕⊕◯◯ - LOW
IMP3	0.36-0.92	0.90-1.00	0.90-1.00*	0.65 (0.41 to 0.90)*	0.04	⊕◯◯◯ - VERY LOW
p16 HD	0.35-1.00	1.00	1.00*	0.62 (0.53 to 0.71)*	0.00	⊕⊕⊕◯ - MODERATE
Mesothelin/SMRP	0.54-0.95	0.48-1.00	0.90 (0.84-0.93)	0.73 (0.68 to 0.77)	0.00	⊕⊕⊕◯ - MODERATE

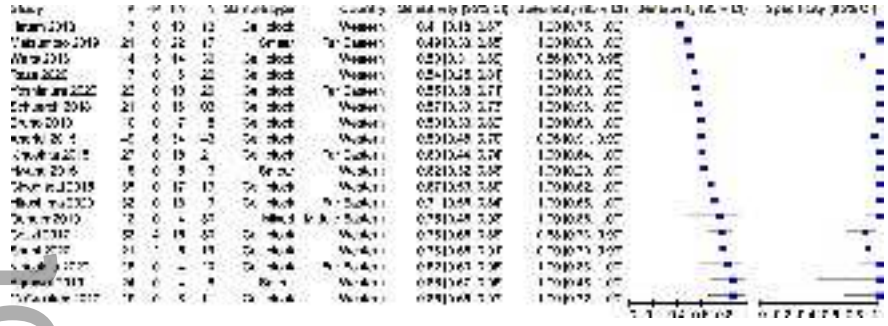
*univariate model



cncy_22509_f1.tif

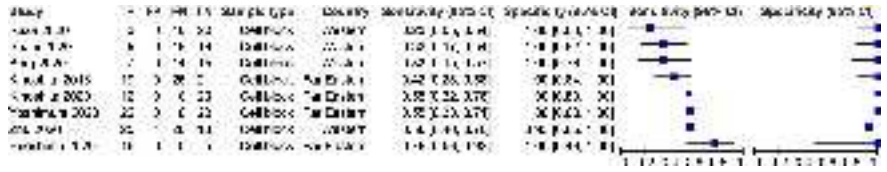


cncy_22509_f2.tiff

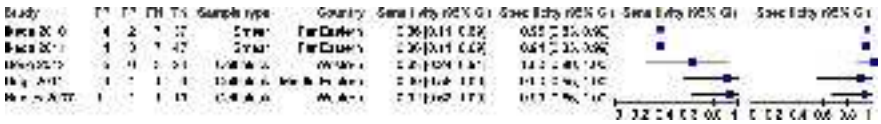


cncy_22509_f3.tif

cncy_22509_f4.tif

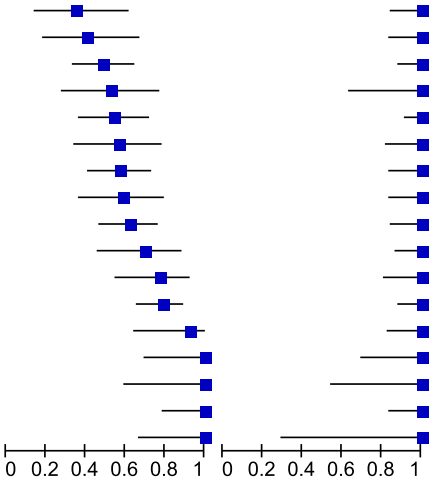


cncy_22509_f5.tif

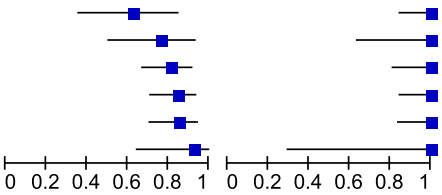


cncy_22509_f6.tif

Study	TP	FP	FN	TN	Sample type	Country	Sensitivity (95% CI)	Specificity (95% CI)
Walts 2016	6	0	11	21	Cell block	Western	0.35 [0.14, 0.62]	1.00 [0.84, 1.00]
Galateau-Salle 2008	7	0	10	20	Cell block	Western	0.41 [0.18, 0.67]	1.00 [0.83, 1.00]
Matsumoto 2019	21	0	22	28	Smear	Far Eastern	0.49 [0.33, 0.65]	1.00 [0.88, 1.00]
Bruno 2019	9	0	8	8	Cell block	Western	0.53 [0.28, 0.77]	1.00 [0.63, 1.00]
Onofre 2008	18	0	15	39	Smear	Western	0.55 [0.36, 0.72]	1.00 [0.91, 1.00]
Flores-Staino 2010	12	0	9	18	Smear	Western	0.57 [0.34, 0.78]	1.00 [0.81, 1.00]
Yoshimura 2020	23	0	17	20	Cell block	Far Eastern	0.57 [0.41, 0.73]	1.00 [0.83, 1.00]
Kinoshita 2020	13	0	9	20	Cell block	Far Eastern	0.59 [0.36, 0.79]	1.00 [0.83, 1.00]
Kinoshita 2018	28	0	17	21	Cell block	Far Eastern	0.62 [0.47, 0.76]	1.00 [0.84, 1.00]
Hida 2015	14	0	6	25	Smear	Far Eastern	0.70 [0.46, 0.88]	1.00 [0.86, 1.00]
Hiroshima 2016	17	0	5	17	Cell block	Far Eastern	0.77 [0.55, 0.92]	1.00 [0.80, 1.00]
Savic 2010	41	0	11	28	Smear	Western	0.79 [0.65, 0.89]	1.00 [0.88, 1.00]
Illei 2003	12	0	1	19	Smear	Western	0.92 [0.64, 1.00]	1.00 [0.82, 1.00]
Nabeshima 2012	10	0	0	10	Smear	Far Eastern	1.00 [0.69, 1.00]	1.00 [0.69, 1.00]
Hamasaki 2012	7	0	0	6	Smear	Far Eastern	1.00 [0.59, 1.00]	1.00 [0.54, 1.00]
Matsumoto 2013	15	0	0	20	Smear	Far Eastern	1.00 [0.78, 1.00]	1.00 [0.83, 1.00]
Hwang 2016	9	0	0	3	Smear	Western	1.00 [0.66, 1.00]	1.00 [0.29, 1.00]



Study	TP	FP	FN	TN	Sample type	Country	Sensitivity (95% CI)	Specificity (95% CI)
Walts 2016	10	0	6	21	Cell block	Western	0.63 [0.35, 0.85]	1.00 [0.84, 1.00]
Bruno 2019	13	0	4	8	Cell block	Western	0.76 [0.50, 0.93]	1.00 [0.63, 1.00]
Matsumoto 2019	35	0	8	17	Smear	Far Eastern	0.81 [0.67, 0.92]	1.00 [0.80, 1.00]
Kinoshita 2018	38	0	7	21	Cell block	Far Eastern	0.84 [0.71, 0.94]	1.00 [0.84, 1.00]
Yoshimura 2020	34	0	6	20	Cell block	Far Eastern	0.85 [0.70, 0.94]	1.00 [0.83, 1.00]
Hwang 2016	12	0	1	3	Smear	Western	0.92 [0.64, 1.00]	1.00 [0.29, 1.00]



Study	TP	FP	FN	TN	Country	Sensitivity (95% CI)	Specificity (95% CI)
Leong 2015	13	0	11	9	Western	0.54 [0.33, 0.74]	1.00 [0.66, 1.00]
Canessa 2013	4	2	3	75	Western	0.57 [0.18, 0.90]	0.97 [0.91, 1.00]
Canessa 2013 (2)	20	1	14	34	Western	0.59 [0.41, 0.75]	0.97 [0.85, 1.00]
Vigani 2011	25	6	13	96	Western	0.66 [0.49, 0.80]	0.94 [0.88, 0.98]
Creaney 2014	122	54	61	592	Western	0.67 [0.59, 0.73]	0.92 [0.89, 0.94]
Creaney 2007	35	2	17	82	Western	0.67 [0.53, 0.80]	0.98 [0.92, 1.00]
Fujimoto 2010	16	8	7	18	Far Eastern	0.70 [0.47, 0.87]	0.69 [0.48, 0.86]
Ferro 2013	30	3	13	33	Western	0.70 [0.54, 0.83]	0.92 [0.78, 0.98]
Creaney 2013	30	3	13	52	Western	0.70 [0.54, 0.83]	0.95 [0.85, 0.99]
Davies 2009	17	2	7	73	Western	0.71 [0.49, 0.87]	0.97 [0.91, 1.00]
Hooper 2012	18	2	7	24	Western	0.72 [0.51, 0.88]	0.92 [0.75, 0.99]
Aleman 2009	13	11	5	10	Western	0.72 [0.47, 0.90]	0.48 [0.26, 0.70]
Battolla 2012	38	9	14	120	Western	0.73 [0.59, 0.84]	0.93 [0.87, 0.97]
Canessa 2012	38	9	14	120	Western	0.73 [0.59, 0.84]	0.93 [0.87, 0.97]
Pass 2008	34	5	11	25	Western	0.76 [0.60, 0.87]	0.83 [0.65, 0.94]
Scherpereel 2006	33	5	10	16	Western	0.77 [0.61, 0.88]	0.76 [0.53, 0.92]
Stockhammer 2020	38	8	10	16	Western	0.79 [0.65, 0.90]	0.67 [0.45, 0.84]
Yamada 2011	36	4	9	20	Far Eastern	0.80 [0.65, 0.90]	0.83 [0.63, 0.95]
Bradley 2013	5	3	1	9	Western	0.83 [0.36, 1.00]	0.75 [0.43, 0.95]
Blanquart 2012	57	4	4	11	Western	0.93 [0.84, 0.98]	0.73 [0.45, 0.92]
Amany 2013	19	2	1	18	Middle Eastern	0.95 [0.75, 1.00]	0.90 [0.68, 0.99]