Clinical identification of malignant pleural effusions

Antonia Marazioti[1,*], Apostolos Voulgaridis[2,*], Ioannis Psallidas[1,3,4,5,*], Anne-Sophie Lamort[6,7,*], Marianthi Iliopoulou[1,*], Anthi C. Krontira[1], Ioannis Lilis[1], Rachelle Asciak[4,5], Nikolaos I. Kanellakis[1,4,5], Najib M. Rahman[4,5], Kyriakos Karkoulias[2], Konstantinos Spiropoulos[2], Jan-Christian Kaiser[8,§], and Georgios T. Stathopoulos[1,6,7,§,#]

[1] Department of Physiology, Faculty of Medicine, University of Patras, Rio, Achaia, 26504, Greece

[2] Department of Pulmonary Medicine, Rio University Hospital, Faculty of Medicine, University of Patras, Rio, Achaia, 26504, Greece

[3] Lungs for Living Research Centre, UCL Respiratory, University College London, London, WC1E 6BT, United Kingdom

[4] Laboratory of Pleural and Lung Cancer Translational Research, Nuffield Department of Medicine, University of Oxford, Oxford, OX3 7FZ, United Kingdom

[5] Oxford Centre for Respiratory Medicine, Churchill Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford, OX3 7LE, UK

[6] Comprehensive Pneumology Center (CPC) and Institute for Lung Biology and Disease (iLBD), Helmholtz Center Munich-German Research Center for Environmental Health (HMGU) and Ludwig-Maximilian-University (LMU), Munich, Bavaria, 81377, Germany

[7] German Center for Lung Research (DZL), Gießen, Hesse, 35392, Germany

[8] Institute of Radiation Medicine, Helmholtz Center Munich-German Research Center for Environmental Health (HMGU), Neuherberg, Bavaria, 85764, Germany.

***** These authors contributed equally to this work as lead authors.

[§] These authors contributed equally to this work as senior authors.

[#] **Corresponding author:** Georgios T. Stathopoulos, Comprehensive Pneumology Center, Max-Lebsche-Platz 31, 1. OG, Munich, 81377, Germany; phone: +498931874846; E-mail: stathopoulos@helmholtz-muenchen.de.

Author Contributions: AM, AV, IP, ASL, MI, ACK, IL, KK, and KS performed and analyzed the clinical study in Patras, Greece; RA, NIK, NMR, and IP performed and analyzed the clinical study in Oxford, UK; JCK did regression and random forest analyses; GTS conceived the main idea and steered the study, performed data analyses and graphic design, and wrote the paper. GTS is the guarantor of the content of the manuscript, including data and analysis. All authors contributed to the conception, design, acquisition, analysis, interpretation, drafting, and revising the work. All authors approved the version to be published and agree to be accountable for the work.

Funding: This work was supported by European Research Council 2010 Starting Independent Investigator (#260524) and 2015 Proof of Concept (#679345) grants, the Graduate College (Graduiertenkolleg, GRK) #2338 of the German Research Society (Deutsche Forschungsgemeinschaft, DFG), the target validation project for pharmaceutical development ALTERNATIVE of the German Ministry for Education and Research (Bundesministerium für Bildung und Forschung, BMBF), and a Translational Research Grant by the German Center for Lung Research (Deutsches Zentrum für Lungenforschung, DZL) (all to GTS). The study sponsors had no role in study design, data collection, analysis, and interpretation, and in writing and submitting the paper for publication.

Running head: The malignant pleural effusion detection score.

Descriptor: 9.31 Pleural Diseases/Mesothelioma

Word count, manuscript: 2,878

Conflict of interest: I.P. is employed as a Global Clinical Head by Astra Zeneca Pharmaceutical in a field that is non-related with the publication. The remaining authors have declared that no conflict of interest exists.

ABSTRACT

Introduction: Pleural effusions frequently signal disseminated cancer. Diagnostic markers of pleural malignancy at presentation that would assess cancer risk and would streamline diagnostic decisions remain unidentified. The objective of the present study was to identify and validate predictors of malignant PE at patient presentation.

Methods: A consecutive cohort of 323 patients with PE from different etiologies was recruited between 2013-2017 and was retrospectively analyzed. Data included history, chest X-ray, and blood/pleural fluid cell counts and biochemistry. Group comparison, receiver-operator characteristics, unsupervised hierarchical clustering, binary logistic regression, and random forests were used to develop the malignant pleural effusion detection (MAPED) score. MAPED was validated in an independent retrospective UK cohort (n = 238).

Results: Five variables showed significant diagnostic power and were incorporated into the 5-point MAPED score. Age > 55 years, effusion size > 50% of the most affected lung field, pleural neutrophil count < 2,500/mm^3, effusion protein > 3.5 g/dL, and effusion lactate dehydrogenase > 250 U/L, each scoring one point, predicted underlying cancer with area under curve = 0.819 ($P < 10^{-15}$) in the derivation cohort and 0.723 ($P = 3 * 10^{-9}$) in the validation cohort. Interestingly, MAPED correctly identified 33/42(79%) of cytology-negative patients that indeed had cancer.

Conclusion: The MAPED score identifies malignant pleural effusions with satisfactory accuracy and can be used complimentary to cytology to streamline diagnostic procedures.

Key words: malignant pleural effusion size; ageing and cancer; lactate dehydrogenase; neutrophil; protein.

Word count, abstract: 222

INTRODUCTION

Pleural effusions (PE) are common conditions that annually affect 1.5 million individuals in the US alone (1). PE are caused by involvement of the pleural space by cancer (malignant PE, MPE) or by non-cancerous processes including infection, inflammation, and deranged Starling or oncotic pressures along juxtapleural blood and lymphatic vessels (benign PE, BPE) (1). Patients with PE face a markedly dichotomous outcome: a diagnosis of MPE portends a median survival of a few months (2–4), while patients with BPE fare significantly better (1). While the time and procedures required for a definitive cell- or tissue-based diagnosis of MPE or an etiologic diagnosis of BPE are substantial, a simple model to predict malignancy that would rapidly inform physicians of the probability of cancer is missing. Even cytological examination of three pleural fluid specimens obtained on consecutive days, considered to be the gold standard in MPE diagnosis together with pleural tissue biopsy, is only successful in two-thirds of MPE cases overall (2–5).

To bridge this gap, we retrospectively evaluated 323 consecutive patients with PE that were admitted to our emergency wards between 2013 and 2017. We collected all clinical, pleural fluid and blood, and chest X-ray data that were available at patient presentation, and performed detailed examination of pleural cells on May-Grünwald-Giemsa-stained cytocentrifugal specimens. MPE-BPE comparisons, receiver-operator characteristics, unsupervised hierarchical clustering, binary logistic regression, and random forests identified five variables that independently predict MPE and that were compiled into the MPE detection (MAPED) model. MAPED displayed area under curve (AUC) = 82% in the derivation cohort and AUC = 72% in a validation cohort of 238 patients with PE from the Oxford Radcliffe Pleural Biobank (ORPB). Interestingly, MAPED performed complementary to cytology.

METHODS

Patients: MAPED and ORPB abided by the Helsinki Declaration, were prospectively approved (University of Patras Ethics Committee #22699/21.11.2013 and South Central Oxford A Ethics Committee #15/SC/0186), and all patients gave written informed consent. All 460 adults with a PE that were admitted to the University Hospital of Patras, Greece, between 21/11/2013–21/11/2017, were evaluated. One hundred and thirty seven patients were excluded due to immediate discharge, previous pleural disease/cancer, and/or missing data. We recorded chest X-ray and blood/PE cell, biochemistry, and pH data. Since routine effusion cell counts were done on smears, we additionally counted 400 cells/patient using May-Gruenwald-Giemsa-stained cytocentrifugal specimens (50,000 cells; 300 g; 10 min; 4 ^{0}C; Cellspin, Tharmac, Wiesbaden, Germany)(6). PE size was defined on the most affected lung field: 1, ≤ 10%; 2, 11-25%; 3, 26-50%; 4, 51-75%; and 5, > 75%. MPE was diagnosed by positive cytology of any of three consecutive daily effusion samples (50-200 mL). Cytology-negative patients were further investigated with thoracoscopy and/or computed tomography-guided biopsy. BPE was diagnosed using established diagnostic criteria (positive pleural fluid smears, cultures, or polymerase chain reaction for common pathogens or Mycobacteria; lymphocytic-predominant exudative effusion with recent tuberculin skin test conversion or conversion within a month after admission; full remission of PE and lung lesions on empiric antibacterial or antituberculous treatment; caseating granulomas in pleural tissue; transthoracic echocardiography-determined ejection fraction < 40% with/without tricuspid regurgitation and/or diastolic dysfunction and/or elevated serum N-terminal pro-B-type natriuretic peptide levels; etc.), according to current guidelines (1-4).

Statistics and analyses: Sample size (*n*) was determined using G*Power (https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-

arbeitspsychologie/gpower)(7). Employing Fischer's exact test, α and β = 0.05, and 1:1 allocation, n = 100-334 was required to detect 20% proportion inequalities between two independent groups, depending on the proportion range (0-100%). Employing Student's t-test, α and β = 0.05, and 1:1 allocation ratio, n = 328 was required to detect effect size d = 0.4 between two independent groups. All data were not normally distributed using Kolmogorov-Smirnov test, hence data are given as frequencies or as median(95% confidence interval, 95%CI). Differences between variables were examined using Fischer's exact, χ^2, or Kolmogorov-Smirnov tests. Probability (P) < 0.05 was considered significant. Bonferroni-corrected P was used for multiple comparisons. Receiver-operator characteristics, unsupervised hierarchical clustering, binary logistic regression using backward Waldman elimination, and random forests were done on Project R* (8) and Prism v8.0 (GraphPad, San Diego, CA). Random forests were grown using R* package randomForest (https://cran.r-project.org/web/packages/randomForest/randomForest.pdf) using log-transformed cell counts, lactate dehydrogenase (LDH), and glucose. The top five variables by the criteria of low P values and mean decreased accuracy were chosen for MAPED, and cut points were determined on partial dependency plots. MAPED was compared with logistic regression and random forest models including all variables, or only variables with lowest Akaike information criteria (AIC) or top variable importance rank (VIMP), using fifty hold-out resampling repeats of 70:30 train/test data sets to calculate AUC and Brier score/skill. For each repeat, a confusion test matrix for P_{MPE} > 0.5 was used to calculate accuracy, sensitivity, and specificity.

Word count, Methods: 499

RESULTS

Distinctive features of malignant versus benign pleural effusions

In total, 323 patients were analyzed in the MAPED study, 189 with BPE and 134 with MPE (Table 1). Most patients with MPE received cytologic diagnoses (n = 92), while 52 received tissue-based, and ten both tissue- and cell-based MPE diagnoses. Cytology was used as the reference standard against which all MAPED variables were tested. Out of the 134 patients with MPE, sixty patients had lung cancer (45%), 30 breast cancer (23%), 21 malignant pleural mesothelioma (16%), 10 gynecological malignancies (7%), four gastrointestinal tumors (3%), five hematological malignancies (4%), and four other cancers (3%), proportions that are in accord with other European studies (2,4). Several differences were identified between patients with BPE and MPE in the 34 different variables examined, using the Bonferroni-adjusted probability threshold of $P < 0.05/34 = 0.0015$ (Figures 1A and 1B, and Table 1). To this end, MPE were more frequently cytology positive and larger in size compared with BPE (Figure 1A). In addition, patients with MPE displayed increased age, PE LDH levels, and PE/blood LDH and protein ratios (Figure 1B). Receiver-operator analyses targeted at MPE identified cytology, PE size, PE neutrophil percentage, and PE-to-blood protein ratio as inputs significantly associated with incipient MPE diagnosis (Figure 1C). Interestingly, PE LDH levels and PE-to-blood LDH and protein ratios represent Light's criteria used to distinguish exudates from transudates (9). In summary, comparative analyses identified variables with some stand-alone predictive power of MPE, which was clearly inferior in comparison to cytology.

Machine learning identifies patient age and effusion size, neutrophil, LDH, and protein contents as the strongest predictors of underlying cancer

Subsequently, binary logistic regression and random forest analyses using MPE as target identified age, PE size, PE neutrophil count, and PE LDH and protein levels as optimal predictors of an incipient MPE diagnosis using the criteria of low probability values and mean decreased accuracy, respectively (Figure 2A). Visual inspection of partial dependency plots defined appropriate cut-offs for accurate distinction of MPE from BPE (Figure 2B). However, Euclidean distancing with agglomeration method ward.D2 (hclust function in R) was incapable of discriminating MPE apart from BPE using unsupervised hierarchical clustering (Figure 3A). Interestingly, all five variables are routinely determined in most hospitals, except PE cytocentrifugal specimen preparation followed by PE differential cell counts, which we routinely implemented (Figure 3B).

Development and validation of the MAPED score

We next incorporated age, PE size, PE neutrophil count, and PE LDH and protein levels into the MAPED score (Figure 4A). MAPED was developed and tested via resampling using the hold-out method based on 50 repeats of train/test data sets split by 70:30 ratios. To assess the predictive power of MAPED in comparison with selected logistic regression and random forest models, benchmark criteria (AUC, Brier score/skill, and AIC) were evaluated. To put the performance of MAPED into perspective, logistic regression models with all available variables or with lowest and top random forest variables were fitted to the data, and random forest models with all available and top variables were considered. Mean and standard deviations from resampling 50 test data sets for the five benchmark criteria are shown in Table 2. These analyses showed that MAPED performed equally or even superior to all random machine learning approaches employed.

In the MAPED cohort, the MAPED score was statistically significantly in fair agreement with cytology results when all patients were examined together, underpinning its value as a

MPE-relevant end-point (Figure 4B). However, MAPED was not in agreement with cytology results when patients with MPE were examined separately (Figure 4C). This fact portrays the potential synergy of the MAPED score with cytology, since MAPED identified 33 out of 42 patients with MPE and negative cytology results (sensitivity = 79%), while it falsely incriminated only 50 patients with BPE out of the 231 with negative cytology as having cancer (specificity = 78%). These findings underscore the potential value of MAPED in the setting of negative cytology results, as well as its potential synergy with cytology in streamlining management: only 9/134(7%) patients with MPE were unidentified by both MAPED and cytology. In addition, MAPED was markedly differently distributed in patients with BPE and MPE, and identified MPE with AUC = 82% in the derivation cohort (Figures 4D-4F).

We finally determined the accuracy of MAPED in the ORPB, one of the few cohorts where PE size was determined and where PE neutrophil data are available, although in a different format compared with MAPED (neutrophil versus lymphocyte predominance is determined in most hospitals worldwide as compared with our quantitative data). Despite these discrepancies, MAPED performed reasonably well in 238 (125 with BPE and 113 with MPE) ORPB patients, correctly predicting MPE with AUC = 72% (Figures 4G and 4H; Table 3). Taking together both cohorts, a MAPED score of > 3 points was determined in $n = 294$ patients and correctly identified 196 of 247 MPE, yielding a sensitivity of 79%, while falsely incriminating 98 of 314 BPE, for a specificity of 69%.

DISCUSSION

MAPED addresses an unmet clinical need: to determine the likelihood of a MPE during initial patient work-up. We show that five variables combined into the MAPED score can predict MPE in 82% of cases in the MAPED derivation cohort, performing almost equally to cytology, and also predicted MPE in 72% of the ORPB validation cohort. We also show that MAPED performs differently than cytology in patients with MPE, and can supplement cytologic investigation in streamlining management. For example, the current clinical practice of repeat cytology and/or tissue biopsy in cytology-negative patients with PE could be improved by prioritizing the 33 MAPED-high cytology-negative patients with MPE and the 50 MAPED-high cytology-negative patients with BPE for intensive cytology (cytospins/cellblocks/immunocytochemistry) and pleural tissue biopsy (n = 83 total patients prioritized) over our 231 total patients investigated. This would come at a cost of only 9 patients (< 3%) with MPE that were missed by both cytology and MAPED, thereby decreasing time to and cost of diagnosis.

The accuracy of MAPED is satisfactory given its simplicity. Another effort to build a similar score included patients with uncertain diagnoses, had multiple primary end-points, and lacked external validation (10). A chest computed tomography (CT)-based score derived from 343 prospectively enrolled patients with PE achieved AUC = 0.919 in discriminating MPE from BPE (11). However, contrast-enhancement and scan reading by two blinded experienced radiologists was required, increasing risk, cost, and time. Despite the above, inter-observer agreement was only 0.55–0.94. To scan our 323 discovery and 238 validation patients assuming a cost of € 200/scan and 0.5 hour physician time required for scan interpretation would cost € 112,200 and 280.5 radiologist hours. We used routine bedside tests to build

MAPED, which performs only slightly inferior to the above-referenced CT score, at zero additional cost and physician time.

The MAPED score components are worth mentioning here. Age is linked with cancer development (12), but its value in prospectively differentiating MPE from BPE has never been exploited, as most studies did not detect age differences between patients with MPE and BPE (1,2,4,13). We did, and although the mean age difference between our patients with BPE and MPE was small (six years), an age cut-off of > 55 years alone could discriminate MPE from BPE with AUC = 0.603. This was not the case in the ORPB validation set, where an age cut-off of > 55 years produced an AUC = 0.526 (P = 0.4925). Notwithstanding population (https://knoema.com/atlas/topics/Demographics/Age/Median-age-of-population?origin=knoema.de; accessed on 08.04.2021) and healthcare accessibility (https://ourworldindata.org/grapher/healthcare-access-and-quality-index; accessed on 08.04.2021) differences between Greece (2015 mean age = 43.4 years and healthcare access & quality index = 87) and the UK (2015 mean age = 40.0 years and healthcare access & quality index = 84.6) that can explain this discrepancy and may necessitate different age cut-offs in different countries, we chose to develop a generally applicable MAPED score and applied it to ORPB patients.

Relative neutrophil predominance in pleural fluid is also a well-known hallmark of infectious BPE due to common pathogens (3), but has not been used as a negative marker of PE malignancy. Interestingly, one important study identified blood neutrophil-to-lymphocyte count ratio as an important determinant of the survival of patients with MPE (2). However, the use of relative pleural neutrophil abundance to rule out MPE is hampered by the common practice of not accurately counting PE cells by most hospitals in the US and Europe. We overcame this by establishing PE differential counts as routine practice for the purposes of

MAPED. The effort was worth spent, since neutrophil counts < 2,500/mm^3 produced a statistically significant AUC = 0.598 (P = 0.0026) in MAPED. Again, this was not the case in ORPB, where pleural neutrophil paucity ≤ 10% produced a non-significant AUC = 0.527 (P = 0.4668). However, neutrophil paucity in ORPB was defined semi-quantitatively on PE smears and not quantitatively on PE cytospins as in MAPED.

Unlike the aforementioned predictors of MPE that failed to perform well in the ORPB validation cohort, PE size > 50% of the most affected lung field, PE protein levels > 3.5 g/dL, and PE LDH levels > 250 U/L did perform excellently in both MAPED and ORPB. In specific, only 27% of BPE but an astonishing 58% of MPE fulfilled the > 50% of the most affected lung field size criterion in ORPB ($P = 6*10^{-5}$; χ^2 test) compared with 15% and 47% in MAPED ($P = 5*10^{-8}$; χ^2 test), respectively. In addition, 60% of BPE and 78% of MPE fulfilled the protein criterion in ORPB (P = .0030; χ^2 test) compared with 68% and 87% in MAPED, respectively ($P = 7*10^{-5}$; χ^2 test); and 65% of BPE and 84% of MPE fulfilled the LDH criterion in ORPB (P = .0007; χ^2 test) compared with 63% and 86% in MAPED, respectively ($P = 6*10^{-6}$; χ^2 test); with the similar results probably owing to more uniform methods of measurement, since pleural protein and LDH levels are established Light's criteria (3, 9). The size, protein, and LDH criteria also produced significant AUC values in ORPB, comparable to those achieved in MAPED (Tables 1 and 2). Although it is well established that MPE pathogenesis includes increased vascular permeability leading to protein-rich exudate (14), pleural fluid-to-blood protein ratio is an exudate criterion according to Light (3, 9), and protein measurements are routine in contemporary hospitals, PE protein levels have never been exploited to diagnose MPE. To this end, PE LDH > 1500 U/L was recently proposed as a poor prognosis marker for MPE (2), and high MPE protein levels were found in a previous study (10), rendering our findings plausible. Massive PE have rarely been studied separately, although they are common with both BPE and MPE (15, 16). In the

largest study looking at PE size, Porcel *et al*. classified 535 patients with BPE and 231 with MPE into three size categories based on posterior-anterior chest X-rays: non-large PE was defined as occupying less than two thirds of the lung field, large as occupying more than that, and massive as occupying the whole lung field (15). Interestingly and in accord with our results, the authors found that 24% of non-large, 49% of large, and 59% of massive PE were malignant ($P < 10^{-5}$; χ^2 test), but this pearl has never been used to estimate MPE risk.

The present study has limitations. First, the general applicability of MAPED is hampered by population, measurement, and practice differences between countries, as well as by divergent prevalence of specific causes of PE. However, MAPED was only loosely fit to the derivation cohort by avoiding weighing, and hence withstood testing in such a remote setting. In addition, the relative prevalence of MPE in MAPED (40% of all PE) was similar to most other published studies from Europe and North America that report values from 30–54% (1,10,16). A second limitation is chest X-ray interpretation, the only non-standardized measure included in MAPED. However, estimating the percentage of a lung field occupied by a PE is an easy task. Third, MAPED cannot be used in outpatients, and patients with previous PE or cancer, by design. Finally, MAPED was not designed to detect MPE that are diagnosed many years past initial PE presentation (17).

In conclusion, the MAPED score correctly classified 412 of 561 (73.44%) pleural effusions examined in two countries, at no additional risk to patients, cost to healthcare systems, and time spent to caring physicians. Pending further validation, MAPED may contribute to improvements in patient management and research design, since it alters the likelihood of MPE at admission as a rule out or rule in score.

REFERENCES

1. Walker SP, Morley AJ, Stadon L, De Fonseka D, Arnold DT, Medford ARL, Maskell NA. Nonmalignant Pleural Effusions: A Prospective Study of 356 Consecutive Unselected Patients. *Chest* 2017;151:1099-105.
2. Clive AO, Kahan BC, Hooper CE, Bhatnagar R, Morley AJ, Zahan-Evans N, Bintcliffe OJ, Boshuizen RC, Fysh ET, Tobin CL, Medford AR, Harvey JE, van den Heuvel MM, Lee YC, Maskell NA. Predicting survival in malignant pleural effusion: development and validation of the LENT prognostic score. *Thorax* 2014;69:1098-104.
3. Light RW. Clinical practice. Pleural effusion. *N Engl J Med* 2002;346:1971-7.
4. Psallidas I, Kanellakis NI, Gerry S, Thézénas ML, Charles PD, Samsonova A, Schiller HB, Fischer R, Asciak R, Hallifax RJ, Mercer R, Dobson M, Dong T, Pavord ID, Collins GS, Kessler BM, Pass HI, Maskell N, Stathopoulos GT, Rahman NM. Development and validation of response markers to predict survival and pleurodesis success in patients with malignant pleural effusion (PROMISE): a multicohort analysis. *Lancet Oncol* 2018;19:930-9.
5. Swiderek J, Morcos S, Donthireddy V, Surapaneni R, Jackson-Thompson V, Schultz L, Kini S, Kvale P. Prospective study to determine the volume of pleural fluid required to diagnose malignancy. *Chest* 2010;137:68-73.
6. Giannou AD, Marazioti A, Spella M, Kanellakis NI, Apostolopoulou H, Psallidas I, Prijovich ZM, Vreka M, Zazara DE, Lilis I, Papaleonidopoulos V, Kairi CA, Patmanidi AL, Giopanou I, Spiropoulou N, Harokopos V, Aidinis V, Spyratos D, Teliousi S, Papadaki H, Taraviras S, Snyder LA, Eickelberg O, Kardamakis D, Iwakura Y, Feyerabend TB, Rodewald HR, Kalomenidis I, Blackwell TS, Agalioti T, Stathopoulos GT. Mast cells mediate malignant pleural effusion formation. *J Clin Invest* 2015;125:2317-34.

7. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007;39:175-191.

8. R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

9. Light RW, Macgregor MI, Luchsinger PC, Ball WC. Pleural effusions: the diagnostic separation of transudates and exudates. *Ann Intern Med* 1972;77:507-13.

10. Porcel JM, Vives M. Differentiating tuberculous from malignant pleural effusions: a scoring model. *Med Sci Monit* 2003;9:CR175-180.

11. Porcel JM, Pardina M, Bielsa S, González A, Light RW. Derivation and validation of a CT scan scoring system for discriminating malignant from benign pleural effusions. *Chest* 2015;147:513-9.

12. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, Abraham J, Adair T, Aggarwal R, Ahn SY, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Barker-Collo S, Bartels DH, Bell ML, Benjamin EJ, Bennett D, Bhalla K, Bikbov B, Bin Abdulhak A, Birbeck G, Blyth F, Bolliger I, Boufous S, Bucello C, Burch M, Burney P, Carapetis J, Chen H, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC, Criqui MH, Cross M, Dabhadkar KC, Dahodwala N, De Leo D, Degenhardt L, Delossantos A, Denenberg J, Des Jarlais DC, Dharmaratne SD, Dorsey ER, Driscoll T, Duber H, Ebel B, Erwin PJ, Espindola P, Ezzati M, Feigin V, Flaxman AD, Forouzanfar MH, Fowkes FG, Franklin R, Fransen M, Freeman MK, Gabriel SE, Gakidou E, Gaspari F, Gillum RF, Gonzalez-Medina D, Halasa YA, Haring D, Harrison JE, Havmoeller R,

Hay RJ, Hoen B, Hotez PJ, Hoy D, Jacobsen KH, James SL, Jasrasaria R, Jayaraman S, Johns N, Karthikeyan G, Kassebaum N, Keren A, Khoo JP, Knowlton LM, Kobusingye O, Koranteng A, Krishnamurthi R, Lipnick M, Lipshultz SE, Ohno SL, Mabweijano J, MacIntyre MF, Mallinger L, March L, Marks GB, Marks R, Matsumori A, Matzopoulos R, Mayosi BM, McAnulty JH, McDermott MM, McGrath J, Mensah GA, Merriman TR, Michaud C, Miller M, Miller TR, Mock C, Mocumbi AO, Mokdad AA, Moran A, Mulholland K, Nair MN, Naldi L, Narayan KM, Nasseri K, Norman P, O'Donnell M, Omer SB, Ortblad K, Osborne R, Ozgediz D, Pahari B, Pandian JD, Rivero AP, Padilla RP, Perez-Ruiz F, Perico N, Phillips D, Pierce K, Pope CA 3rd, Porrini E, Pourmalek F, Raju M, Ranganathan D, Rehm JT, Rein DB, Remuzzi G, Rivara FP, Roberts T, De León FR, Rosenfeld LC, Rushton L, Sacco RL, Salomon JA, Sampson U, Sanman E, Schwebel DC, Segui-Gomez M, Shepard DS, Singh D, Singleton J, Sliwa K, Smith E, Steer A, Taylor JA, Thomas B, Tleyjeh IM, Towbin JA, Truelsen T, Undurraga EA, Venketasubramanian N, Vijayakumar L, Vos T, Wagner GR, Wang M, Wang W, Watt K, Weinstock MA, Weintraub R, Wilkinson JD, Woolf AD, Wulf S, Yeh PH, Yip P, Zabetian A, Zheng ZJ, Lopez AD, Murray CJ, AlMazroa MA, Memish ZA. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;380:2095-128.

13. Taghizadeh N, Fortin M, Tremblay A. US Hospitalizations for Malignant Pleural Effusions: Data From the 2012 National Inpatient Sample. *Chest* 2017;151:845-54.

14. Stathopoulos GT, Kalomenidis I. Malignant pleural effusion: tumor-host interactions unleashed. *Am J Respir Crit Care Med* 2012;186:487-92.

15. Porcel JM, Vives M. Etiology and pleural fluid characteristics of large and massive effusions. *Chest* 2003;124:978-83.

16. Ryu J-S, Ryu HJ, Lee S-N, Memon A, Lee S-K, Nam H-S, Kim HJ, Lee KH, Cho JH, Hwang SS. Prognostic impact of minimal pleural effusion in non-small-cell lung cancer. *J Clin Oncol* 2014;32:960-7.

17. Koegelenberg CFN, Irusen EM, von Groote-Bidlingmaier F, Bruwer JW, Batubara EMA, Diacon AH. The utility of ultrasound-guided thoracentesis and pleural biopsy in undiagnosed pleural exudates. *Thorax* 2015;70:995-7.

FIGURE LEGENDS

Figure 1. Variables significantly different between benign (BPE) and malignant (MPE) pleural effusions (PE) in the MPE detection (MAPED) study. (A) Frequency distributions of PE cytology and size stratified by PE diagnosis. Data are presented as patient numbers (*n*) with Fischer's exact (left) and χ^2 (right) test probabilities (*P*). **(B)** Continuous numerical variables stratified by diagnosis. Shown are patient numbers (*n*), raw data (dots), rotated kernel density distributions (violins), median (dashed lines), quartiles (dotted lines), and Kolmogorov-Smirnov test probabilities (*P*). LDH, lactate dehydrogenase. **(C)** Shown are significant receiver-operator characteristics (curves), areas under curve (AUC) with 95% confidence intervals (95%CI), and probabilities (*P*).

Figure 2. Machine learning identifies patient age, pleural effusion (PE) size, as well as PE neutrophil, protein, and lactate dehydrogenase (LDH) content as important predictors of malignant PE (MPE) in the MAPED study. (A) Variable importance (VIMP) plot. Data are presented as estimates (circles), cut-offs (dashed lines), and variables selected or not for inclusion into MAPED (color code). **(B)** Partial dependency plots from random forests (RF) in comparison to linear binary LR. Data are presented as probability of MPE versus benign PE (BPE) by each predictor.

Figure 3. Unsupervised hierarchical clustering alone cannot distinguish benign from malignant pleural effusions (PE). **(A)** Heatmap shows unsupervised hierarchical clustering of *n* = 323 patients by patient age, PE size, as well as PE neutrophil (NΦ), protein, and lactate dehydrogenase (LDH) content. Each row represents one patient and each column one marker. **(B)** Representative PE cytocentrifugal specimens (cytospins) with differential cell counts. Arrows indicate neutrophils. MAPED, malignant pleural effusion detection study and score.

Figure 4. The malignant pleural effusion detection (MAPED) score and its performance in the Greek discovery and a UK validation cohort. (A) Graphic representation of the

MAPED score and its constituents. **(B, C)** Cross-tabulations of MAPED score by pleural effusion (PE) cytology results in all patients (B) and in patients with malignant PE (MPE; C). Shown are patient numbers (n), coefficients of agreement (κ), and Fischer's exact probabilities (*P*). **(D)** Data summary of MAPED score by pleural effusion (PE) diagnosis (MPE versus benign PE, BPE). Shown are patient numbers (*n*), raw data (circles), rotated kernel density distributions (violins), median (dashed lines), quartiles (dotted lines), and Kolmogorov-Smirnov test probability (*P*). **(E, F)** Probability of MPE by MAPED score (E) and receiver-operator curve of MAPED targeting MPE diagnosis (F) in the Greek MAPED derivation cohort. Data in (E) are presented as fractions (columns), color-coded patient numbers, and probability (*P*), χ^2 test. Data in (F) are presented as receiver-operator characteristic (curve) with area under curve (AUC), 95% confidence interval (95% CI), and probability (*P*). *n*, sample size. **(G, H)** Probability of MPE by MAPED score (G) and receiver-operator curve of MAPED targeting MPE diagnosis (H) in the UK Oxford Radcliffe Pleural Biobank (ORPB) validation cohort. Data in (G) are presented as fractions (columns), color-coded patient numbers, and probability (*P*), χ^2 test. Data in (H) are presented as receiver-operator characteristic (curve) with area under curve (AUC), 95% confidence interval (95% CI), and probability (*P*). *n*, sample size.

TABLES

Table 1. Summary of malignancy of pleural effusion detection (MAPED) study patient data at entry stratified by primary outcome.

	BPE[a]	**MPE**[b]	***P***[c]	**ROC AUC (95%CI)**[d]	**ROC *P***[e]
n[f]	189	134	-	-	-
Cytology (negative/positive)[f]	189/0	42/92	$< 10^{-15}$	.843(.794-.893)	$< 10^{-15}$
Tumor tissue of origin (none/lung/breast/pleura/female GU/gut/blood/other) [f,g]	189/0/0/0/0/0/0/0	0/60/30/21/10/4/5/4	$< 10^{-15}$	1.000(1.000-1.000)	$< 10^{-15}$
Age (years)[h]	64(61–67)	70(68–72)	.0014	.603(.542-.664)	.0016
Sex (male/female)[f]	93/96	54/80	.1132	.545(.481-.608)	.1725
Smoking (never/former/current)[f]	62/78/49	39/59/36	.7748	.517(.453-.581)	.6056
PE side (right/left/bilateral)[f,i]	96/70/23	80/48/6	.0420	.519(.456-.582)	.5632
PE size score (1/2/3/4/5)[f,i,j]	13/66/81/21/8	6/25/40/41/22	$5*10^{-8}$	.672(.611-.734)	10^{-7}
PE red blood cells (10^6/mm^3)[h,i]	2.7(1.8-4.0)	6.1(4.7-10.5)	.0017	.587(.524-.651)	.0075
PE nucleated cells (10^6/mm^3)[h,i]	1.8(1.5-2.1)	1.6(1.3-1.9)	.0856	.540(.478-.603)	.2185
White blood cells (10^6/mm^3)[h]	8.5(8.0-9.5)	8.9(8.5-9.7)	.1196	.538(.475-.601)	.2471
PE large mononuclear cells (%)[f,i]	45(41-55)	47(44-54)	.2784	.536(.473-.599)	.2690
PE neutrophils (%)[h,i]	7(4-10)	3(2-4)	.0059	.623(.562-.684)	.0002
PE small mononuclear (lymphoid) cells (%)[h,i]	19(14-27)	28(24-42)	.0245	.602(.540-.664)	.0018
PE eosinophils (%)[h,i]	0(0-1)	0(0-1)	.9525	.510(.447-.574)	.7510
PE large mononuclear cells (10^6/mm^3)[h,i]	.74(.58-.90)	.76(.51-.94)	.5061	.509(.446-.573)	.7767
PE neutrophils (10^6/mm^3)[h,i]	.09(.07-.13)	.04(.02-.05)	.0064	.598(.537-.660)	.0026
PE small mononuclear (lymphoid) cells (10^6/mm^3)[h,i]	.28(.19-.36)	.46(.28-.56)	.0306	.573(.511-.636)	.0247
PE eosinophils (10^3/mm^3)[h,i]	0(0-0)	0(0-8)	.5919	.512(.449-.576)	.7120
Blood mononuclear cells (%)[h]	8(7-8)	8(7-8)	.9913	.518(.454-.582)	.5793

Blood neutrophils (%)[h]	70(67-73)	70(69-71)	.3771	.521(.458-.584)	.5180
Blood lymphocytes (%)[h]	18(16-20)	19(17-22)	.5740	.548(.485-.611)	.1427
Blood eosinophils (%)[h]	2(1-2)	2(1-2)	.0982	.557(.494-.620)	.0802
Blood mononuclear cells (10^6/mm^3)[h]	.65(.61-.71)	.63(.60-.71)	.6443	.505(.441-.568)	.8865
Blood neutrophils (10^6/mm^3)[h]	5.9(5.4-6.3)	6.1(5.7-6.7)	.1748	.527(.464-.590)	.4130
Blood lymphocytes (10^6/mm^3)[h]	1.5(1.4-1.7)	1.7(1.5-1.8)	.0184	.577(.514-.639)	.0190
Blood eosinophils (10^6/mm^3)[h]	.15(.12-.17)	.14(.11-.15)	.3549	.542(.479-.606)	.1946
PE LDH (U/L)[h,i,k]	325(271-450)	505(436-627)	.0002	.582(.520-.644)	.0119
Blood LDH (U/L)[h,k]	338(313-369)	366(307-392)	.6532	.518(.453-.583)	.5755
PE/Blood LDH ratio[h,i,k]	1.1(0.9-1.3)	1.5(1.2-1.7)	.0002	.575(.513-.637)	.0220
PE protein (g/dL)[h,i]	4.2(4.1-4.6)	4.6(4.5-4.8)	.0055	.586(.524-.647)	.0085
Blood protein (g/dL)[h]	6.9(6.7-7.0)	6.7(6.5-6.8)	.1386	.536(.473-.600)	.2643
PE/Blood protein ratio[h,i]	.65(.62-.67)	.68(.66-.70)	.0004	.606(.545-.667)	.0011
PE glucose (mg/dL)[h,i]	101(95-106)	103(98-107)	.6938	.520(.456-.584)	.5414
Blood glucose (mg/dL)[h]	102(98-106)	103(100-108)	.0265	.522(.459-.584)	.5099
PE/Blood glucose ratio[h,i]	1.00(.95-1.00)	1.00(.96-1.00)	.6384	.504(.440-.567)	.9119
PE pH[h,i]	7.38(7.37-7.40)	7.38(7.36-7.40)	.4945	.524(.460-.588)	.4677

[a] BPE, benign pleural effusion.
[b] MPE, malignant pleural effusion.
[c] *P*, probability, χ^2, Fischer's exact, or Kolmogorov-Smirnov test. Threshold of significance by Bonferroni correction $P = 0.05/34$ comparisons = 0.0015.
[d] ROC AUC (95%CI), receiver-operator characteristic area under curve (95% confidence interval).
[e] ROC *P*, receiver-operator characteristic probability. Threshold of significance by Bonferroni correction $P = 0.05/34$ comparisons = 0.0015.
[f] Data presented as number of patients (*n*).
[g] GU, genitourinary.
[h] Data presented as median(95% confidence interval).
[i] PE, pleural effusion
[j] Pleural effusion size score as % of largest lung field affected as assessed by chest X-ray, ultrasound, or computed tomography: 1, ≤ 10%; 2, 11-25%; 3, 26-50%; 4, 51-75%; 5, > 75%.
[k] LDH, lactate dehydrogenase.

Table 2. Power for MPE prediction from resampling 50 test data sets for all covariables (allCV)[a], variables pertaining to the lowest Akaike information criterion (AIC)[b], the five benchmark components of the malignant pleural effusion detection (MAPED) score (topCV), and the MAPED score per se by logistic regression (LR) and random forests (RF) analyses. Shown are performance measures on 70% training and 30% test data sets from 50 bootstraps. In bold font are shown the outperformers.

Model	**AUC[c,d]**	**Accuracy[d]**	**Sensitivity[d]**	**Specificity[d]**	**Brier score[d]**	**Brier skill[e]**
LR, allCV	.765±.040	.715± .035	.620±.073	.784±.055	.204±.024	-.261
RF, allCV	.808±.035	.731±.036	.543±.073	**.866±.052**	.183±.010	-.134
LR, low AIC	.818±.037	.756±.037	.685±.071	.807±.048	.174±.020	-.075
LR, topCV	.791±.043	.752±.040	.635±.076	.835±.051	.183±.018	-.134
RF, topCV	.811±.037	.727±.032	.628±.064	.798±.049	.174±.018	-0.080
LR, MAPED	**.822±.033**	**.774±.032**	**.815±.050**	.744±.046	**.161±.017**	**Ref[f]**

[a] Covariables entered in raw form: smoking status, PE side, PE size, age, sex, and PE and blood differential cell percentages; and in log-transformed form: white blood and PE nucleated cell counts, PE red blood cell count, PE and blood differential cell counts, PE pH, PE and blood lactate dehydrogenase (LDH), protein, and glucose levels.

[b] Covariables entered in raw form: PE size, age, PE side, and PE large mononuclear, neutrophil, and eosinophil cell percentages; and in log-transformed form: PE large mononuclear cell count, PE pH, LDH, and protein levels, and blood protein levels.

[c] AUC, area under curve.

[d] Data presented as mean±SD.

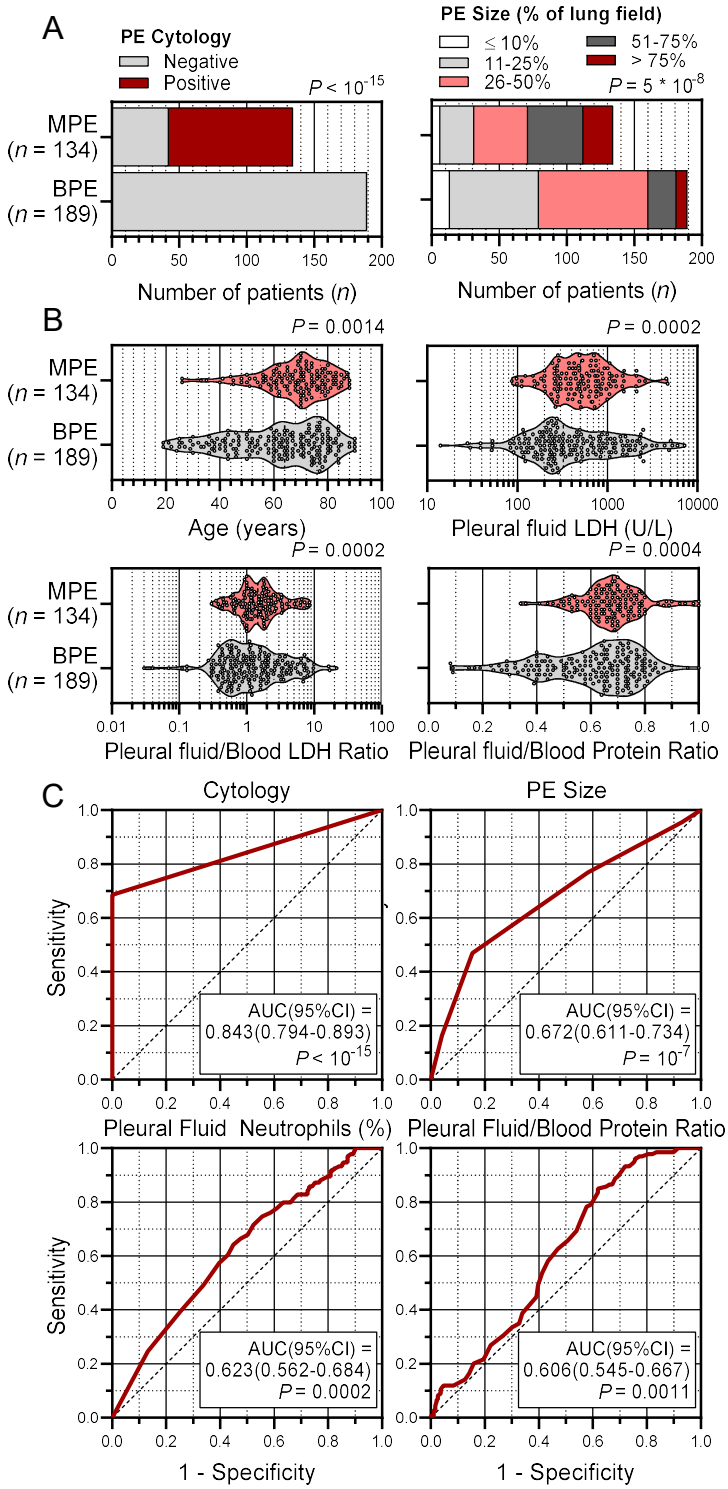
[e] A Brier skill score has a range of -∞ to 1. Negative values mean that the forecast is less accurate than a standard forecast. < 0 = lesser skill compared to the reference forecast. 0 = no skill compared to the reference forecast. 1 = perfect skill compared to the reference forecast.

[f] Ref, Reference.

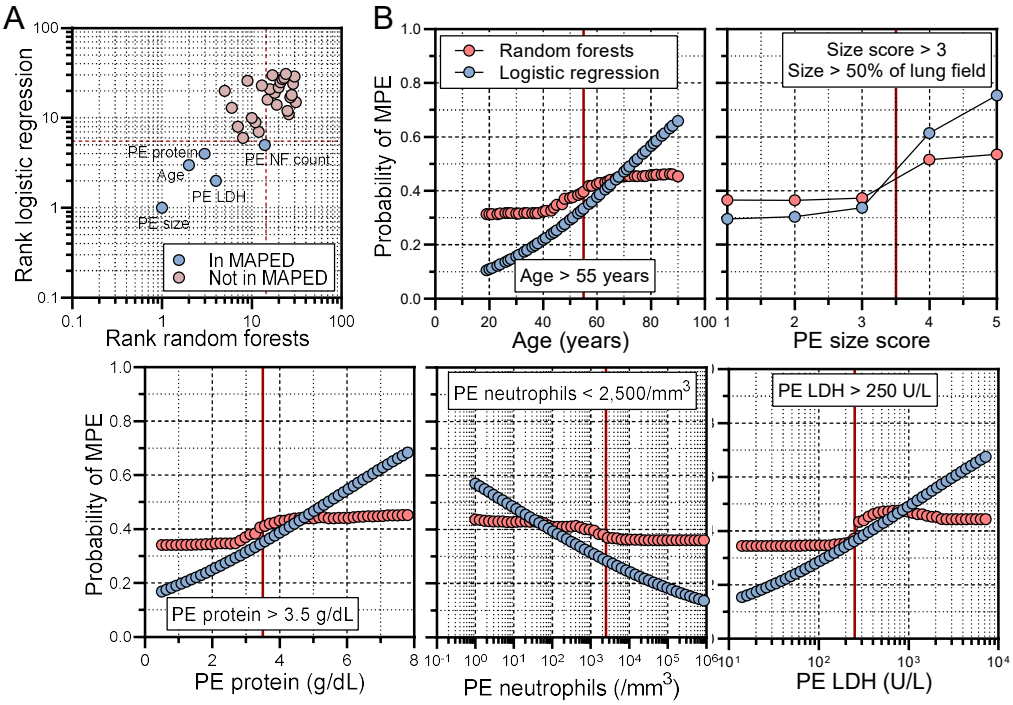
Table 3. Summary of Oxford Radcliffe Pleural Biobank (ORPB) study patient data stratified by primary outcome.

	BPE[a]	**MPE**[b]	***P***[c]	**ROC AUC (95%CI)**[d]	**ROC *P***[e]
n[f]	125	113	-	-	-
Age (years)[g]	75(70-78)	72(71-76)	.4715	.526(.452-.599)	.4925
PE side (unilateral/bilateral)[f,h]	114/11	103/10	.9893	.500(.427-.574)	.9947
PE size score (1/2/3/4/5)[f,h,i]	7/8/76/14/20	5/5/38/36/29	$6*10^{-5}$	.634(.562-.705)	.0004
PE neutrophils > 10% (no/yes)[f,h]	106/19	102/11	.2059	.527(.454-.601)	.4668
PE LDH (U/L)[g.h,k]	356(333-459)	642(571-802)	.0002	.673(.604-.743)	$7*10^{-6}$
PE protein (g/dL)[g,h]	3.8(3.6-4.1)	4.3(4.1-4.4)	.0111	.619(.548-.690)	.0015

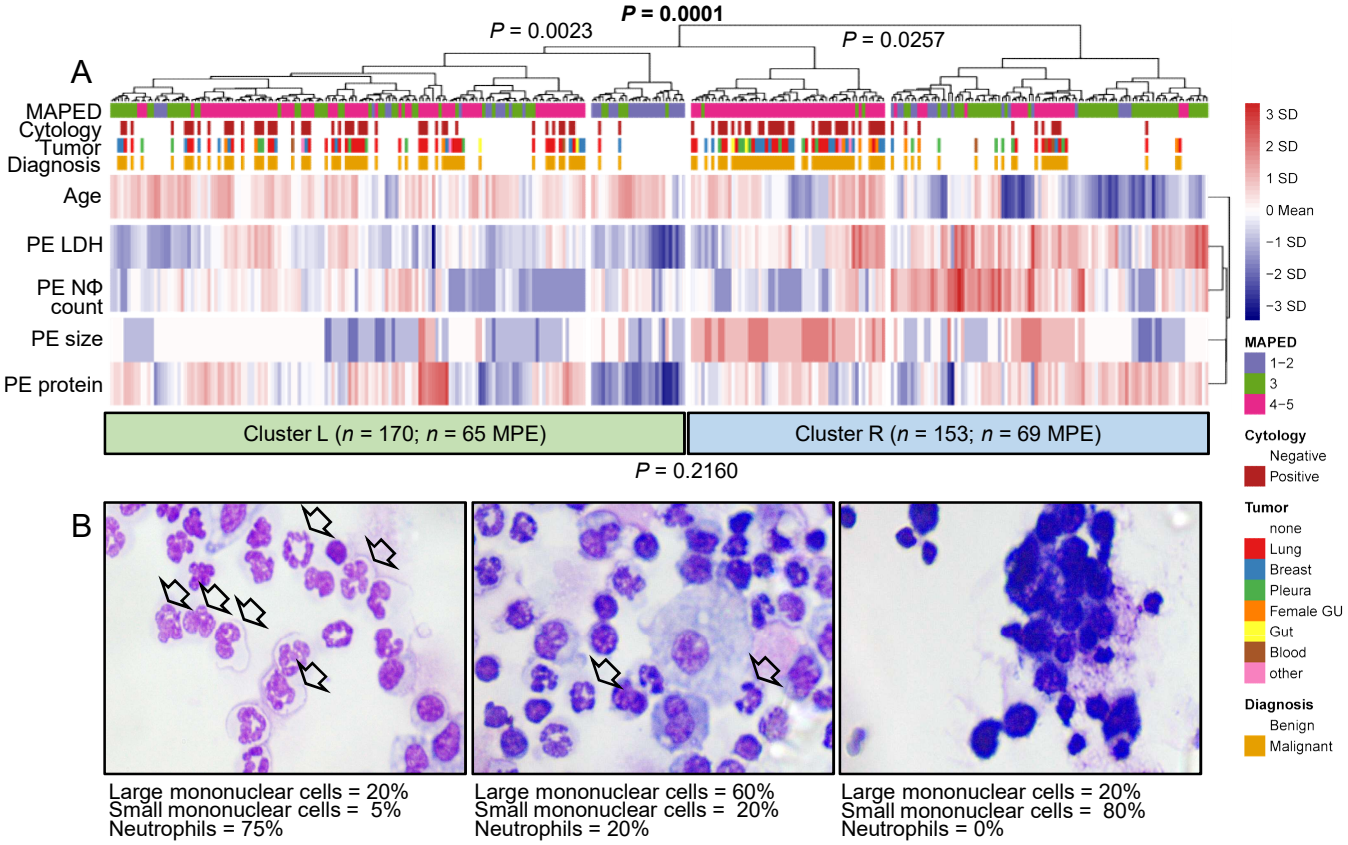
[a] BPE, benign pleural effusion.
[b] MPE, malignant pleural effusion.
[c] *P*, probability, χ^2, Fischer's exact, or Kolmogorov-Smirnov test. Threshold of significance by Bonferroni correction $P = 0.05/6$ comparisons = 0.0083.
[d] ROC AUC (95%CI), receiver-operator characteristic area under curve (95% confidence interval).
[e] ROC *P*, receiver-operator characteristic probability. Threshold of significance by Bonferroni correction $P = 0.05/6$ comparisons = 0.0083.
[f] Data presented as number of patients (*n*).
[g] Data presented as median(95% confidence interval).
[h] PE, pleural effusion
[i] Pleural effusion size score as % of largest lung field affected as assessed by chest X-ray, ultrasound, or computed tomography: 1, ≤ 10%; 2, 11-25%; 3, 26-50%; 4, 51-75%; 5, > 75%.
[k] LDH, lactate dehydrogenase.



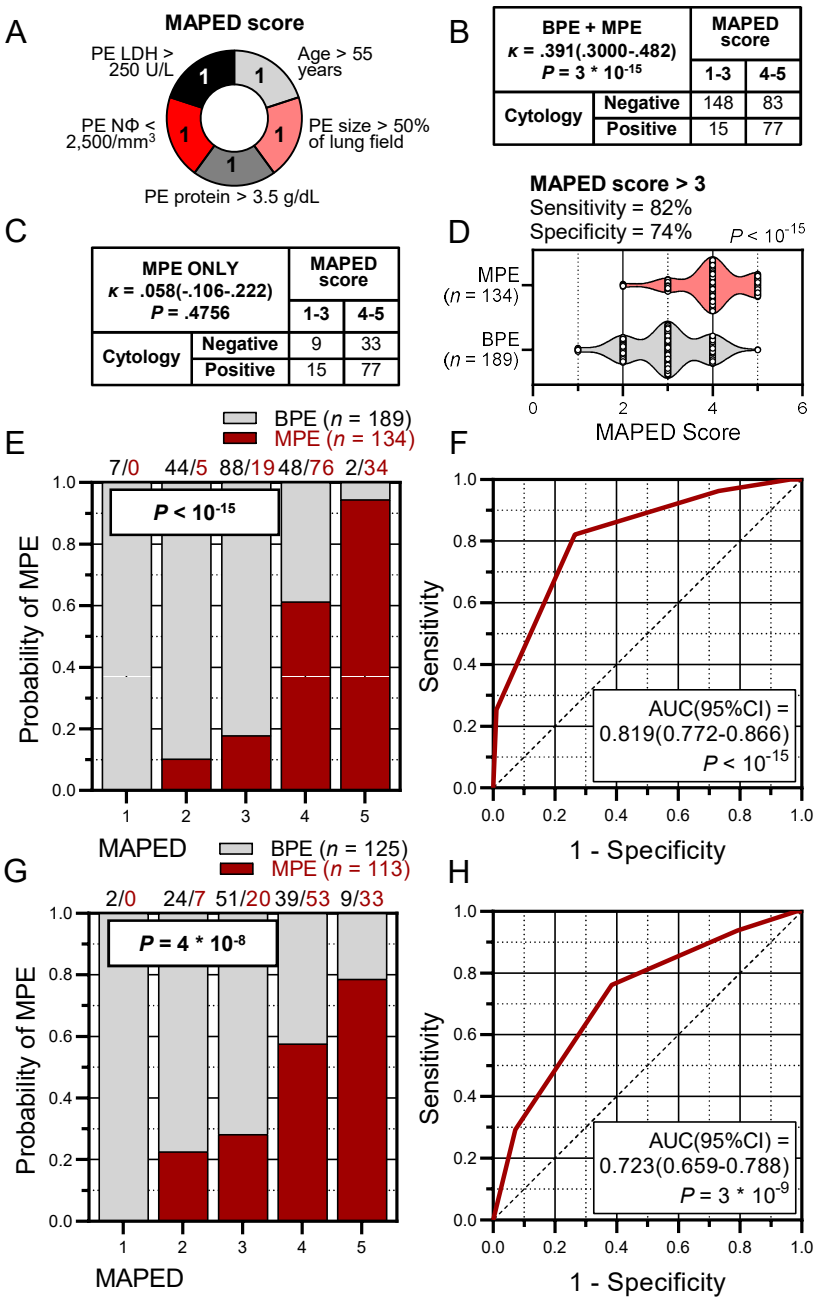
Marazioti *et al*. **Figure 1**



Marazioti *et al*. **Figure 2**



Marazioti *et al*. **Figure 3**



Marazioti *et al*. **Figure 4**