ORIGINAL ARTICLE

Predicting survival in malignant pleural effusion: development and validation of the LENT prognostic score

Amelia O Clive,[1,2] Brennan C Kahan,[3] Clare E Hooper,[1,2] Rahul Bhatnagar,[1,2] Anna J Morley,[2] Natalie Zahan-Evans,[2] Oliver J Bintcliffe,[2] Rogier C Boshuizen,[4] Edward T H Fysh,[5,6] Claire L Tobin,[5] Andrew R L Medford,[2] John E Harvey,[2] Michel M van den Heuvel,[4] Y C Gary Lee,[5,6] Nick A Maskell[1,2]

Linked

CrossMark

ABSTRACT

Background Malignant pleural effusion (MPE) causes debilitating breathlessness and predicting survival is challenging. This study aimed to obtain contemporary data on survival by underlying tumour type in patients with MPE, identify prognostic indicators of overall survival and develop and validate a prognostic scoring system.

Methods Three large international cohorts of patients with MPE were used to calculate survival by cell type (univariable Cox model). The prognostic value of 14 predefined variables was evaluated in the most complete data set (multivariable Cox model). A clinical prognostic scoring system was then developed and validated.

Results Based on the results of the international data and the multivariable survival analysis, the LENT prognostic score (pleural fluid lactate dehydrogenase, Eastern Cooperative Oncology Group (ECOG) performance score (PS), neutrophil-to-lymphocyte ratio and tumour type) was developed and subsequently validated using an independent data set. Risk stratifying patients into low-risk, moderate-risk and high-risk groups gave median (IQR) survivals of 319 days (228–549; n=43), 130 days (47–467; n=129) and 44 days (22–77; n=31), respectively. Only 65% (20/31) of patients with a high-risk LENT score survived 1 month from diagnosis and just 3% (1/31) survived 6 months. Analysis of the area under the receiver operating curve revealed the LENT score to be superior at predicting survival compared with ECOG PS at 1 month (0.77 vs 0.66, p<0.01), 3 months (0.84 vs 0.75, p<0.01) and 6 months (0.85 vs 0.76, p<0.01).

Conclusions The LENT scoring system is the first validated prognostic score in MPE, which predicts survival with significantly better accuracy than ECOG PS alone. This may aid clinical decision making in this diverse patient population.

Key messages

What is the key question?
► What are the key prognostic indicators in malignant pleural effusion (MPE) and could they be used to develop a clinical risk score to help predict survival and guide management?

What is the bottom line?
► Survival of patients with MPE varies greatly and given the invasive nature of the treatment options available, better prognostication at presentation could help select the optimal management strategy for individual patients.

Why read on?
► The LENT score is the first validated risk stratification system to predict survival in MPE, calculated on the basis of pleural fluid lactate dehydrogenase, Eastern Cooperative Oncology Group performance score, serum neutrophil-to-lymphocyte ratio and tumour type.

INTRODUCTION

Malignant pleural effusion (MPE) is common, with an estimated annual incidence of 150 000 in the USA alone and given the year-on-year increase in new cancer diagnoses, the incidence is set to rise.[1 2] MPE represents advanced malignant disease and current guidelines quote median survivals of between 3 and 12 months.[3] Pleural and oncological treatment options are expanding and hence more accurate prognostication at presentation may help individualise treatment strategies.[3 4] Treatments may cause morbidity and necessitate an inpatient stay; hence, identifying patients with the poorest prognosis may help to minimise discomfort and inconvenience at the end of their lives.

A number of factors may help predict survival of patients with malignant pleural disease, including tumour characteristics, extent of disease, comorbidities and the composition of the effusion.[5–8] In addition, inflammation-based scoring systems have been associated with overall prognosis in a number of cancer types, including mesothelioma.[9–12] However, until now, no validated score exists to risk stratify unselected patients with MPE.

The goals of this study were to (a) obtain contemporary data on overall survival and survival according to underlying cell type for patients with MPE and (b) identify other prognostic indicators of overall survival in unselected patients with MPE, with the aim of developing and validating a prognostic scoring system to help inform clinical management decisions.

METHODS

Three prospectively collected databases from the UK ('UK Cohort 1'), Australia ('Australian Cohort') and The Netherlands ('Dutch Cohort') were used to identify patients with MPE, who had been followed up for a minimum of 12 months or until death (see online supplementary appendix). The three cohorts comprised of patients presenting with their first episode of MPE, which may have represented a new cancer diagnosis or recurrence/progression of previously diagnosed malignancy.

A separate 'UK Cohort 2' was extracted from the UK database using patients with MPE entered at a later time point to that of 'UK Cohort 1' (see online supplementary appendix). The patients in 'UK Cohort 2' had been followed up for a minimum of 6 months or until death.

Local ethical approval was obtained for each of the participating centres and written informed consent was obtained from the studies' participants.

Survival time was calculated from the date of diagnosis of MPE to death (see online supplementary appendix). Subjects were censored at the time of data extraction or loss to follow-up.

Patient data from the 'UK Cohort 1', 'Dutch Cohort' and 'Australian Cohort' were amalgamated to evaluate the effect of underlying cell type on survival, using a univariable Cox model and Kaplan–Meier survival curves.

Fourteen predefined variables, recorded at the time of presentation with MPE, were then evaluated in 'UK Cohort 1' to ascertain their impact on patient survival. This cohort was chosen as it had prospectively collected data for the key potential prognostic variables, minimal missing data and represented routine, unselected practice in a general hospital making the findings widely applicable. Pleural fluid vascular endothelial growth factor levels and serum N-terminal prohormone of brain natriuretic peptide (NT-proBNP) levels were quantified using stored, frozen samples (see online supplementary appendix). The NLR was calculated by dividing the blood neutrophil count by the blood lymphocyte count.

An a priori statistical analysis plan was approved (see online supplementary appendix). A multivariable, proportional hazard Cox model with backward selection was performed to determine which factors were independently predictive of survival. Variables with a p value >0.1 were removed from the model. Eighteen (8.1%) patients in 'UK Cohort 1' had missing data for at least one variable; therefore, multiple imputation with 10 imputed data sets was used to account for this, using chained equations.[13] Fractional polynomials were used for continuous predictors.[14]

A prognostic scoring system was derived from a subset of variables, selected based on the data from the international cohort, the strength of the hazard ratios (HRs) from the multivariable analysis of 'UK Cohort 1' and their clinical availability.

The prognostic score ('the LENT score') was evaluated using patients with complete data for the four selected variables from 'UK Cohort 1'. Overall survival according to the LENT score was assessed using a Cox model. The area under the curve (AUC) for 1-month, 3-month and 6-month mortality and the C statistic were also calculated. The C statistic is a summary measure of a model's accuracy, with values ranging from zero to one.[15] Sensitivity analyses were performed to evaluate the reliability of the score in the following subgroups: (a) only patients with histo-cytologically confirmed malignancy, (b) only patients with malignant cell types other than mesothelioma, (c) patients with missing data included as both best-case and worst-case scenarios and (d) only patients with a new

diagnosis of malignant disease, rather than recurrent or progressive disease.

The LENT score was validated using patients with no missing data for the four variables included in the LENT score from 'UK Cohort 2'. Overall survival was assessed using a Cox model, and the C statistic and AUC for mortality at 1, 3 and 6 months were calculated. Sensitivity analyses were conducted, assuming both best-case and worst-case scenarios for patients with missing baseline variables.

All analyses were performed using Stata software V.12.1 (Texas, USA).

RESULTS

Obtaining contemporary survival data according to underlying malignant cell type

Data were obtained on 789 patients (221 patients from 'UK Cohort 1', 253 from the 'Australian Cohort' and 315 from the 'Dutch Cohort') (see table 1). The mean age was 66.3 (SD 12.9) years, 53.6% were men and 58.4% of effusions were right sided. The 'Dutch Cohort' contained more women, had a lower average age and a shorter median survival than the other cohorts, possibly reflecting the smaller proportion of mesothelioma cases. The 'Australian Cohort' had a higher proportion of histo-cytologically proven MPE.

Survival analysis of the combined cohorts showed a large spread of survival times, ranging from median survivals of under 50 days in the urological cancer, sarcoma and melanoma groups, compared with 339 days for those patients with mesothelioma (see table 2 and figure 1). The lung cancer group had a median survival of only 74 days.

Evaluating other potential prognostic factors

Two hundred and twenty-one patients with a final diagnosis of MPE were included in 'UK Cohort 1'. One hundred and fifty-five of the 221 patients in 'UK Cohort 1' had newly diagnosed malignancy as opposed to 66 patients where the MPE represented recurrence or progression of previously diagnosed cancer.

Based on univariable analysis of the imputed data set, 11 of the 14 preselected variables showed a statistically significant effect on survival (see online supplementary appendix). However, multivariable analysis using backward selection identified effusion size, Eastern Cooperative Oncology Group (ECOG) performance score (PS), pleural fluid lactate dehydrogenase (LDH), serum neutrophil-to-lymphocyte ratio (NLR), serum NT-proBNP and malignant cell type as the only variables to maintain independent associations with survival at our predefined cut-off of p<0.1 (see online supplementary appendix).

Development of the LENT score

On the basis of clinical applicability and the results of the multivariable analysis, we selected four variables (pleural fluid LDH, ECOG PS, NLR and tumour type) to include in a predictive model, 'the LENT score'. Table 3 describes the scoring system, which results in a score range of 0–7. For ease of interpretation, patients can be divided into low-risk (score 0–1), moderate-risk (score 2–4) and high-risk (score 5–7) prognostic groups based on their LENT score calculated at the time of presentation with MPE. The Kaplan–Meier survival curves are shown in figure 2.

Data from 203 patients were used to calculate the summary statistics for the LENT score (18 were excluded as they had incomplete data). The Harrell's C index for the model was 0.71. Patients with moderate-risk and high-risk LENT scores had HRs (95% CI) for mortality of 1.49 (1.03 to 2.15) and

Table 1 Baseline characteristics of the UK, Australian and Dutch cohorts

	UK Cohort 1 n=221	Australian Cohort n=253	Dutch Cohort n=315	UK Cohort 2 n=83
Age, median (IQR)	74 (67–80)	69 (63–78)	60 (53–66)	71 (61–76)
% Men	62.0	65.6	38.1	60.2
% Right sided	57.0	61.0	57.3	63.9
Mode of diagnosis, n (%)				
Pleural cytology, flow cytometry or pleural biopsy	150 (68)	205 (81)	213 (68)	54 (65)
Otherwise unexplained effusion with confirmed malignancy elsewhere	54 (24)	48 (19)	102 (32)	25 (30)
Radiological evidence of malignancy with no histo-cytological proof	17 (8)	0	0	4 (5)
Cell type, n (%)				
Mesothelioma	58 (26)	96 (38)	16 (5)	17 (20)
Lung cancer	66 (30)	72 (28)	77 (24)	33 (40)
Gynaecological cancer	20 (9)	17 (7)	22 (7)	4 (5)
Breast cancer	26 (12)	25 (10)	89 (28)	11 (13)
Gastrointestinal cancer	10 (5)	8 (3)	43 (14)	3 (4)
Haematological malignancy	13 (6)	15 (6)	7 (2)	7 (8)
Other	28 (13)	20 (8)	61 (19)	8 (10)
Overall median survival, days (95% CI)	168 (108 to 228)	205 (167 to 238)	84 (72 to 115)	193 (97 to 332)

5.97 (3.58 to 9.97) compared with those with a low-risk LENT score (see figure 2).

The patients with a low-risk LENT score had a median survival of 319 days (IQR 228–549; n=43); 100%, 98% and 86% survived to 1, 3 and 6 months, respectively. Those with a moderate-risk LENT score had a median survival of 130 days (47–467; n=129) and 81% survived to 1 month, 59% to 3 months and 47% survived to 6 months. This compares with those with a high-risk LENT score who had a median survival of only 44 days (22–77, n=31) and their chances of surviving 1, 3 and 6 months were 65%, 13% and 3%, respectively (see figures 3 and 4).

The receiver operating characteristic (ROC) analysis for mortality at 1, 3 and 6 months using the LENT score gave significantly higher AUC values than ECOG performance status alone. The AUC at 1 month was 0.77 (95% CI 0.70 to 0.84) for LENT compared with 0.66 (0.57 to 0.75) for ECOG PS (p=0.005). At 3 months, the AUC for LENT was 0.84 (0.79 to 0.89) compared with 0.75 (0.69 to 0.81) for ECOG PS (p<0.001). At 6 months, the AUC for LENT was 0.85 (0.80 to 0.90) compared with 0.76 (0.70 to 0.82) for ECOG PS (p=0.001) (see figure 4).

All sensitivity analyses performed using 'UK Cohort 1' showed the results to be robust (see online supplementary appendix).

Validation of the LENT score
Seventy-six of the 83 patients in 'UK Cohort 2' were used for the validation (two patients had missing data for PS, three for fluid LDH, one for tumour type and one for NLR).

Multivariable analysis evaluating the four components of the LENT score using 'UK Cohort 2' showed that ECOG performance and NLR were independent predictors of mortality at the predefined cut-off of p<0.1 (see online supplementary appendix).

Survival analysis of 'UK Cohort 2' revealed that the LENT risk groups had similar HR for mortality as 'UK Cohort 1' (see figure 2). Harrell's C statistic for 'UK Cohort 2' was better than that of 'UK Cohort 1' (0.75 and 0.71, respectively) and the

Table 2 Median survival according to cell type for the UK, Australian and Dutch cohorts combined

Cell type	Median survival in days (95% CI)	n
Mesothelioma	339 (267 to 422)	170
Haematological malignancy	218 (160 to 484)	35
Gynaecological malignancy	203 (97 to 279)	59
Breast cancer	192 (133 to 271)	140
Renal cell carcinoma	114 (33 to 334)	22
Adenocarcinoma of unknown primary	87 (13 to 286)	11
Lung cancer	74 (60 to 92)	215
Other	71 (46 to 102)	33
Gastrointestinal cancer	61 (44 to 73)	61
Sarcoma	44 (19 to 76)	12
Melanoma	43 (23 to 72)	23
Urological cancer (bladder, prostate, testis, penile)	33 (22 to 168)	8
Overall	136 (119 to 167)	789

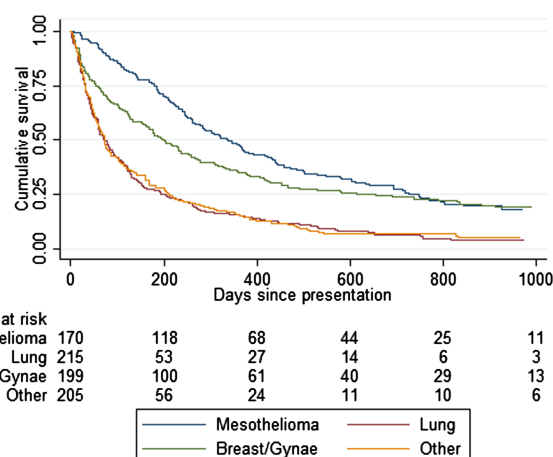| Number at risk						
Mesothelioma	170	118	68	44	25	11
Lung	215	53	27	14	6	3
Breast/Gynae	199	100	61	40	29	13
Other	205	56	24	11	10	6

Figure 1 Kaplan–Meier survival curves according to cell type for the UK, Australian and Dutch cohorts combined.

Table 3 The **LENT** score calculation

	Variable	Score
L	LDH level in pleural fluid (IU/L)	
	<1500	0
	>1500	1
E	ECOG PS	
	0	0
	1	1
	2	2
	3–4	3
N	NLR	
	<9	0
	>9	1
T	Tumour type	
	Lowest risk tumour types	0
	▶ Mesothelioma	
	▶ Haematological malignancy	
	Moderate risk tumour types	1
	▶ Breast cancer	
	▶ Gynaecological cancer	
	▶ Renal cell carcinoma	
	Highest risk tumour types	2
	▶ Lung cancer	
	▶ Other tumour types	

Risk categories	Total score
Low risk	0–1
Moderate risk	2–4
High risk	5–7

ECOG PS, Eastern Cooperative Oncology Group performance score; LDH, lactate dehydrogenase; NLR, neutrophil-to-lymphocyte ratio.

proportion of patients surviving to 1, 3 and 6 months was comparable with that of 'UK Cohort 1' (see figure 3).

The ROC analysis for 'UK Cohort 2' again gave higher AUC values for the LENT score than for ECOG PS. At 1 month, the AUC for the LENT score was 0.78 (95% CI 0.64 to 0.92) compared with 0.69 (0.46 to 0.92) for ECOG PS (p=0.360). At 3 months, the AUCs for the LENT score and ECOG PS were 0.76 (0.65 to 0.86) and 0.62 (0.49 to 0.74), respectively (p=0.002). The AUC at 6 months for the LENT score was also significantly better than ECOG PS (0.81 (0.72 to 0.90) and 0.65 (0.54 to 0.77), respectively) (p<0.001; see figure 4).

A sensitivity analysis was performed, which revealed no important differences when including patients with missing data as either the best-case or worst-case scenarios (see online supplementary appendix).

DISCUSSION

This study is the largest prospectively collected series in the literature to evaluate prognostic factors in patients with MPE. We have developed the first validated risk stratification system for this diverse patient group (the LENT score), which provides valuable additional prognostic information for doctors and patients and performs significantly better than ECOG performance status alone. This information is of direct clinical relevance in tailoring treatment strategies according to an individual's predicted survival.

Our prospective, contemporary and internationally relevant data support data from previous smaller series highlighting the wide range of survivals in this population according to malignant cell type.[6 16] The sizeable difference in median survival between, for example, mesothelioma (339 days) and lung cancer (74 days) is likely to significantly alter the choice of management in these patients. Interestingly, the overall median survival for our entire cohort is not dissimilar to older studies, highlighting that despite oncological advances, outcomes have not substantially improved for these patients.[6 16]

Multivariable analysis of 'UK Cohort 1' highlighted potentially important roles for local and systemic inflammation in predicting survival, independent of the more predictable associations with cell type and ECOG PS.[7 8]

Our data support previous work indicating that high pleural fluid LDH levels (reflecting localised, acute inflammation, necrosis and cell death within the pleural cavity) are indicative of a poor prognosis in MPE.[6 17] Acute, neutrophilic tumour infiltration has also been shown to confer a worse prognosis in mesothelioma.[18]

The highly significant association between NLR and survival supports data from other cancer types, including mesothelioma, suggesting that systemic inflammation is also an important prognostic factor.[11 12 19] This is consistent with the growing
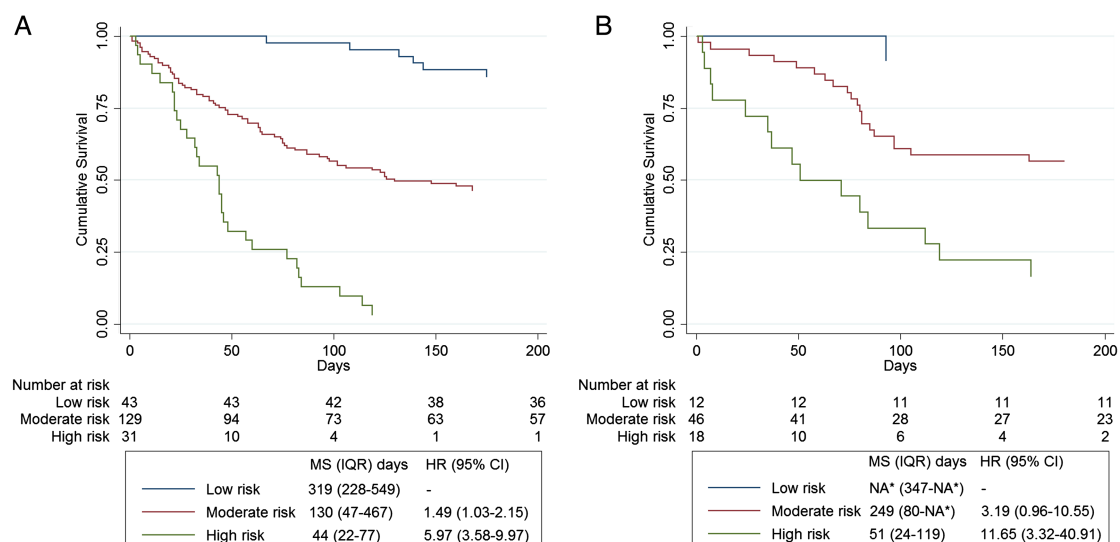
Figure 2 Survival curves according to the LENT score. (A) 'UK Cohort 1' and (B) 'UK Cohort 2'. MS, median survival; IQR, interquartile range; HR, hazard ratio; 95% CI, 95% confidence interval; NA*, Unable to provide data as insufficient patients have died.

Figure 3 Proportion of patients surviving to 1, 3 and 6 months according to low-risk, moderate-risk and high-risk LENT scores. (A) 'UK Cohort 1' and (B) 'UK Cohort 2'.
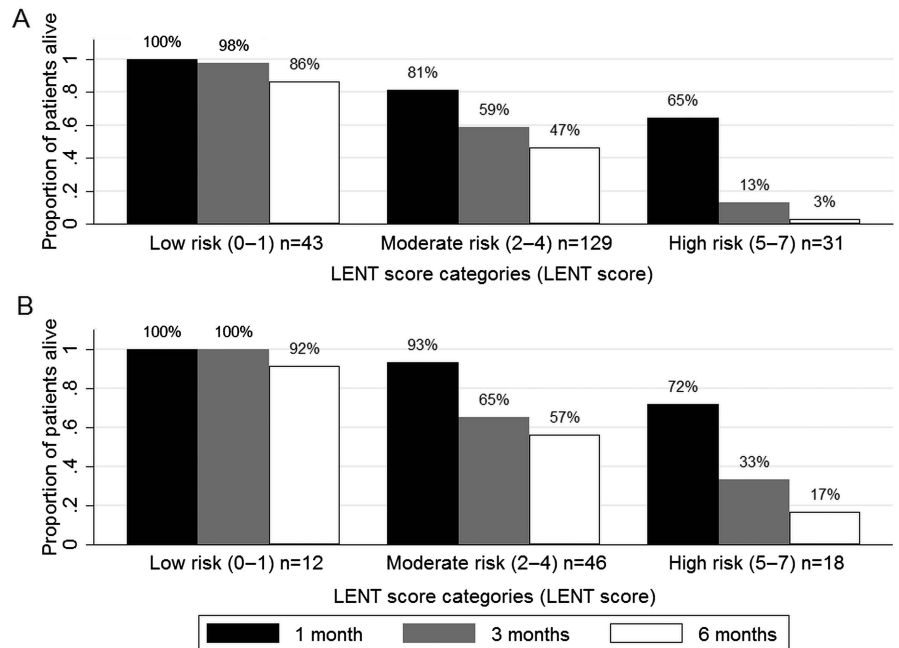
evidence that systemic inflammation may contribute to both morbidity (in the form of constitutional symptoms) and mortality in cancer and may represent a common pathway in the fatal progression of malignancy.[10] Along with the presence of underlying comorbidities (such as cardiac or renal failure), the systemic inflammatory response may also contribute to the poorer prognosis we have identified in those patients with higher NT-proBNP levels through increasing cardiac work.

NLR is simply calculated from a full blood count and hence is cheap and universally available, making it an attractive potential biomarker. Further work is necessary to understand the

mechanisms behind its prognostic value and why it appears to prognosticate better than other inflammatory markers.

Data from two, large randomised controlled trials (RCTs) evaluating treatment strategies in MPE suggest that clinicians are inaccurate at predicting survival in this population. In the TIME2 trial, 36/106 participants (34%) died within 3 months of trial entry (despite an exclusion criteria for trial entry of predicted survival of <3 months) and 58/340 (17%) of patients died within 30 days of trial entry to the Dresler *et al* study (who only included patients with a predicted survival of >2 months).[20 21] This highlights that clinical judgement alone is
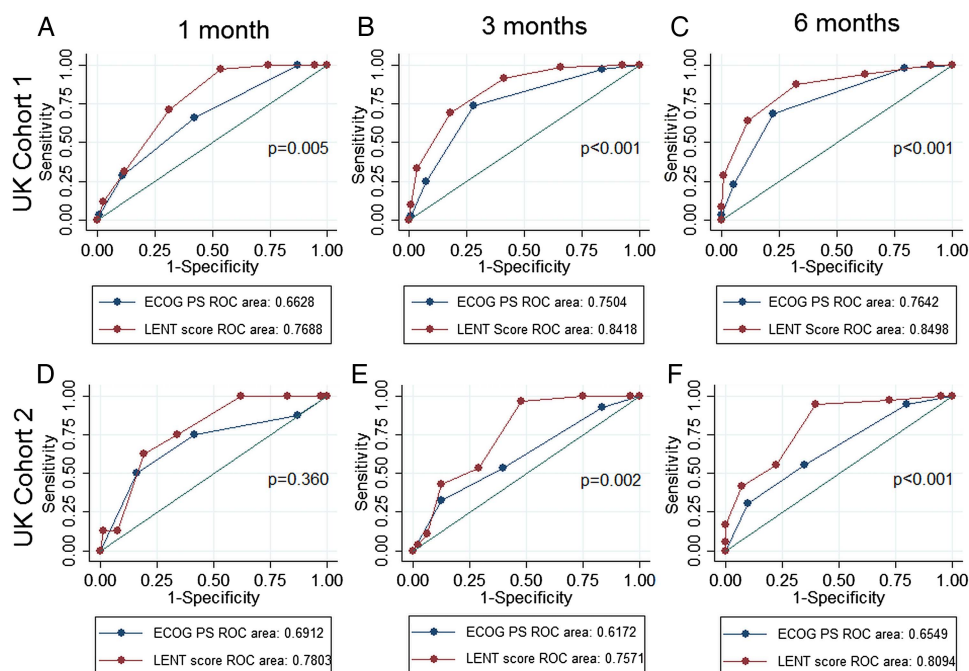
Figure 4 Receiver operating characteristic (ROC) curve analysis for the LENT score and Eastern Cooperative Oncology Group (ECOG) performance score (PS) for the outcome of mortality. (A) 'UK Cohort 1' at 1 month. (B) 'UK Cohort 1' at 3 months. (C) 'UK Cohort 1' at 6 months. (D) 'UK Cohort 2' at 1 month. (E) 'UK Cohort 2' at 3 months. (F) 'UK Cohort 2' at 6 months.

imprecise at estimating patient survival. In addition, the ECOG performance status of patients with MPE may vary as the effusion is managed and hence solely relying only on this to predict prognosis may be suboptimal.

Attempts have been made in a number of cancer types to develop prognostic scoring systems for this purpose.[9] [10] In mesothelioma, the European Organisation for Research and Treatment of Cancer and Cancer and Leukaemia Group B scores have been developed, but no scoring system exists for all-comers with MPE.[22] [23]

In order for the LENT score to be clinically useful and easy to calculate, we selected four variables to include in the score, which are readily available in the clinical setting and had highly significant associations with survival from the multivariable analysis. We have shown the LENT score to be particularly effective at identifying patients with the worst prognosis and that it is significantly better than ECOG PS alone at identifying these patients.

This is important to help patients and clinicians select an optimal management strategy. Patients with a particularly poor prognosis may wish to minimise time spent in hospital by choosing an indwelling pleural catheter or therapeutic pleural aspiration over attempted pleurodesis to manage their effusion.[20] Previous studies report the median length of hospital stay of between 4 and 6 days for inpatient talc pleurodesis, which for patients with a high LENT score would represent a substantial proportion of their remaining life, given their median survival in this study of only 44 days.[20] [24]

Additionally, a robust clinical risk score is likely to be invaluable in future RCTs in this area. The LENT score could be used as an eligibility criterion to select patients likely to survive sufficient time for the primary endpoint analysis, thereby ensuring studies remain suitably powered. It could also be used to balance randomisation groups in the clinical trial setting as a stratification/minimisation criterion.[25]

The strengths of this study include its large sample size and the prospective nature of the case series. By including all patients in the survival analysis and accounting for missing data using multiple imputation or sensitivity analyses, we have ensured that all available data are used and the potential of missing important associations by using complete case data is minimised. We have also validated our risk model using an independent data set, which adds strength to our findings. The LENT scoring system is less subjective than PS alone and relies on readily available tests, allowing it to be rapidly calculated without additional costs. Additionally, as the LENT score has been developed using patients presenting with their first episode of MPE, regardless of previous cancer treatment, it is widely applicable and relevant in the clinical setting.

There are some limitations to the study. The LENT score has been developed using patients with robust diagnoses of MPE and hence may not be generalised to all patients with clinically suspected MPE or para-malignant effusions. We lacked data on the extent that patients had been pretreated for their malignancy and only had data for 'UK Cohort 1' on whether their effusion represented a new malignant diagnosis or recurrent/progressive disease. These factors may impact on survival, particularly in certain cell types such as breast cancer. Due to the diverse staging systems for the different pleural tumours, and the lack of a universal biomarker, it was not possible to include the extent of disease in our analysis. This would be interesting to evaluate in future studies as the associations between NLR and fluid LDH with survival, for example, may be surrogates for tumour burden.

The predominance of mesothelioma cases in this study reflects the high incidence of this tumour in the UK and Western Australia. However, sensitivity analysis found the predictive value of the LENT score remained robust when these patients were excluded, suggesting it is also a valid tool in populations with a lower incidence of mesothelioma. Additionally, given the size of the 'UK Cohort 2', further endorsement with another prospectively collected data set would be beneficial before the LENT score is introduced to standard clinical practice.

This is the largest prospectively collected series in the literature evaluating prognostic factors in MPE. It highlights the disparate survival prospects in this population and identifies markers of local and systemic inflammation as key prognostic indicators. The LENT score is the first validated risk stratification system in MPE and is significantly better than ECOG PS at predicting survival. It is a robust, easy to calculate, clinically relevant prognostic score, which may help discussions with patients about prognosis and thereby guide the selection of appropriate treatment pathways.

REFERENCES

1 The American Thoracic Society. Management of malignant pleural effusions. *Am J Respir Crit Care Med* 2000;162:1987–2001.

2 Cancer Statistics Registrations, England (Series MB1): Office of National Statistics, Stationary Office, 2010.

3 Roberts ME, Neville E, Berrisford RG, *et al*. Management of a malignant pleural effusion: British Thoracic Society Pleural Disease Guideline 2010. *Thorax* 2010;65 (Suppl 2):ii32–40.

4 Tan C, Sedrakyan A, Browne J, *et al*. The evidence on the effectiveness of management for malignant pleural effusion: a systematic review. *Eur J Cardiothorac Surg* 2006;29:829–38.

5 Pilling JE, Dusmet ME, Ladas G, *et al*. Prognostic factors for survival after surgical palliation of malignant pleural effusion. *J Thorac Oncol* 2010;5:1544–50.

6 Bielsa S, Salud A, Martinez M, *et al*. Prognostic significance of pleural fluid data in patients with malignant effusion. *Eur J Intern Med* 2008;19:334–9.

7 Ozyurtkan MO, Balci AE, Cakmak M. Predictors of mortality within three months in the patients with malignant pleural effusion. *Eur J Intern Med* 2010;21:30–4.

8 Burrows CM, Mathews WC, Colt HG. Predicting survival in patients with recurrent symptomatic malignant pleural effusions: an assessment of the prognostic values of physiologic, morphologic, and quality of life measures of extent of disease. *Chest* 2000;117:73–8.

9 Pinato DJ, Mauri FA, Ramakrishnan R, *et al*. Inflammation-based prognostic indices in malignant pleural mesothelioma. *J Thorac Oncol* 2012;7:587–94.

10 Proctor MJ, Morrison DS, Talwar D, *et al*. A comparison of inflammation-based prognostic scores in patients with cancer. A Glasgow Inflammation Outcome Study. *Eur J Cancer* 2011;47:2633–41.

11 Kao SC, Pavlakis N, Harvie R, *et al*. High blood neutrophil-to-lymphocyte ratio is an indicator of poor prognosis in malignant mesothelioma patients undergoing systemic therapy. *Clin Cancer Res* 2010;16:5805–13.

12 Anevlavis S, Kouliatsis G, Sotiriou I, *et al*. Prognostic factors in patients presenting with pleural effusion revealing malignancy. *Respiration* 2014;87:311–6.

13 White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;30:377–99.

14 Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. *Int J Epidemiol* 1999;28:964–74.

15 Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361–87.

16 Heffner JE, Nietert PJ, Barbieri C. Pleural fluid ph as a predictor of survival for patients with malignant pleural effusions. *Chest* 2000;117:79–86.

17 Martinez-Moragon E, Aparicio J, Sanchis J, et al. Malignant pleural effusion: prognostic factors for survival and response to chemical pleurodesis in a series of 120 cases. *Respiration* 1998;65:108–13.

18 Suzuki K, Kadota K, Sima CS, et al. Chronic inflammation in tumor stroma is an independent predictor of prolonged survival in epithelioid malignant pleural mesothelioma patients. *Cancer Immunol Immunother* 2011;60:1721–8.

19 Li MX, Liu XM, Zhang XF, et al. Prognostic role of neutrophil-to-lymphocyte ratio in colorectal cancer: a systematic review and meta-analysis. *Int J Cancer* 2014;134:2403–13.

20 Davies HE, Mishra EK, Kahan BC, et al. Effect of an indwelling pleural catheter vs chest tube and talc pleurodesis for relieving dyspnea in patients with malignant pleural effusion: the TIME2 randomized controlled trial. *JAMA* 2012;307:2383–9.

21 Dresler CM, Olak J, Herndon JE II, et al. Phase III intergroup study of talc poudrage vs talc slurry sclerosis for malignant pleural effusion. *Chest* 2005;127:909–15.

22 Curran D, Sahmoud T, Therasse P, et al. Prognostic factors in patients with pleural mesothelioma: the European Organization for Research and Treatment of Cancer experience. *J Clin Oncol* 1998;16:145–52.

23 Herndon JE, Green MR, Chahinian AP, et al. Factors predictive of survival among 337 patients with mesothelioma treated between 1984 and 1994 by the Cancer and Leukemia Group B. *Chest* 1998;113:723–31.

24 Putnam JB Jr, Walsh GL, Swisher SG, et al. Outpatient management of malignant pleural effusion by a chronic indwelling pleural catheter. *Ann Thorac Surg* 2000;69:369–75.

25 Kahan BC, Morris TP. Reporting and analysis of trials using stratified randomisation in leading medical journals: review and reanalysis. *BMJ* 2012;345:e5840.

On-Line Appendix

Additional methodology regarding the patient cohorts

Three different databases were used to identify the patients used in this study:

(a) A database of consecutive patients presenting to North Bristol NHS Trust, UK was used to identify patients with an ultimate diagnosis of MPE, as confirmed independently by 2 respiratory consultants. From this database, two separate cohorts were extracted:
   a. The 'UK Cohort 1' comprised of 221 patients with MPE entered into the database between March 2008 and April 2012. All patients were followed up for a minimum of 12 months or until death.
   b. The 'UK Cohort 2' comprised of 83 patients with MPE entered into the database between May 2012 and April 2013. Patients were followed up for a minimum of 6 months or until death.
(b) The 'Dutch cohort' was collected from the Netherlands Cancer Institute between June 2006 and May 2012. It comprised of 315 consecutive patients with proven cancer and MPE.
(c) The 'Australian cohort' was obtained from the 'Longitudinal follow up of management of MPE in Western Australia' Study conducted between August 2009 and October 2011, and comprised of 253 patients. This study prospectively recorded patient data from the time of diagnosis of MPE for a minimum of 6 months or until death.

For the purposes of survival analysis, the date of diagnosis was defined using the date pleural malignancy was confirmed by pleural biopsy or cytology. In the absence of histo-cytological confirmation from the pleura, the date of diagnosis was defined by the date that the consistent radiology was performed.

197 of 221 patients in 'UK Cohort 1' died during the study follow up. The remaining 24 patients were censored after a median follow up of 774 days (IQR 473-1058).

52 of 83 patients in 'UK Cohort 2' died during the study follow up. The remaining 31 patients were censored after a median follow up of 220 days (IQR 284-388).

286 of 315 patients in 'Dutch cohort' died during the study follow up. The remaining 29 patients were censored after a median follow up of 941 days (IQR 314-1043).

195 of 253 patients in the 'Australian Cohort' died during the study follow up. The remaining 58 patients were censored after a median follow up of 708 days (IQR 476-922).

Additional methodology regarding sample processing for VEGF and NT-proBNP

For the 'Bristol Cohort' plasma (collected in EDTA), serum and pleural fluid (collected in EDTA and serum gel tubes), were collected at the time of study entry. These samples were centrifuged at 1000G for 20mins and the supernatant frozen at -70°C for the purpose of subsequent analysis.

Pleural fluid VEGF quantification

Commercially available enzyme linked immunoassay (ELISA) kits were used to quantify VEGF levels in the pleural fluid samples (Quantikine © for human VEGF, R&D Systems, Abingdon, UK). Samples were processed in duplicate according to manufacturer's guidelines. The accurate measurement range as determined by the standard solutions is 31.2-2000pg/mL and samples were diluted if necessary.

Serum NT-proBNP quantification

Serum samples were utilized to measure NT-proBNP levels. This was performed using an automated, quantitative Enzyme Linked Fluorescent Assay (ELFA) technique, according to the manufacturers' guidelines (VIDAS ® NT-proBNP, bioMerieux SA, France). The accurate measurement range of the assay is 20-25,000 pg/mL.

A priori statistical analysis plan

Background

The 'pleural investigation trial' is a prospective observational series consisting of consecutive patients who presented to North Bristol NHS Trust with an undiagnosed pleural effusion. Written informed consent was taken from all patients included in the study, following which blood and pleural fluid samples (at the time of trial entry) were stored as per the study protocol. All patients were followed up for a minimum of 6 months or until death. A final diagnosis as to the cause of their effusion was documented.

This series includes the 221 patients who were enrolled into the study between March 2008 and April 2012, who had a final diagnosis of malignant pleural effusion (MPE). The diagnosis of MPE was confirmed independently by 2 respiratory consultants, with at least one of the following criteria having to be present:

- Histocytological proof of pleural malignancy on a pleural biopsy or cytology;
- Histological proof of malignancy elsewhere with a presumed malignant effusion;
- Radiological diagnosis of pleural malignancy without histocytological proof

Analysis Plan

Descriptive statistics will be used to summarize patient characteristics and clinical data. The appropriate statistical tests will be selected to compare the groups and look for correlations depending on whether the data is parametric or non-parametric. For the purposes of analysis of continuous variables, a linear relationship will be assumed in the first instance.

The overall survival of patients will be calculated in days from the date of trial entry to the date of death. Initially, each of the following potential prognostic factors will be evaluated individually using a univariable Cox model and Kaplan Meier Survival Curves to determine if they are independently associated with survival.

Patient/Tumor variables:

Variable	Variable type	% missing data
Age	Continuous	0
Baseline ECOG performance Score	Categorical (0-4)	5
Malignant cell type	Categorical (lung, mesothelioma, breast/ gynecological, other)	0

Comorbidities that may exacerbate MPE:

Variable	Variable type	% missing data
Baseline serum albumin level	Continuous	2
Baseline eGFR	Analyze as categories: <15, 15-29, 30-59, 60-89, >90	0.5
Serum BNP level at baseline	Continuous	10

Inflammation based variables:

Variable	Variable type	% missing data

Neutrophil to lymphocyte ratio in the blood	Continuous	1
Modified Glasgow Prognostic Score (mGPS)	Categorical (0-2)	5
Platelet to lymphocyte ratio (PLR)	Continuous	1
Baseline serum CRP	Continuous	3.1
Baseline pleural fluid VEGF	Continuous	8.5

Pleural Fluid characteristics:

Variable	Variable type	% missing data
Size of the effusion at baseline on CXR	Categorical (1-5)	4
Pleural fluid LDH	Continuous	3.1
Pleural fluid pH and glucose	Categorical (High or Low) (Patients will be divided into 2 categories according to their pleural fluid pH (<7.3; ≥7.3). If no pleural fluid pH value is available, but the pleural fluid glucose has been measured, those with a pleural fluid glucose <3.3mmol/l will be included in the 'low' group and those with a value ≥3.3mmol/l will be included in the 'high' group)	10

In order to determine which variables are still prognostic, even when used in conjunction with other variables, multivariable analysis will be used to assess their adjusted contribution on survival. Only variables with ≤10% missing data will be included in the multivariable analysis. Multiple imputation will be used to account for missing baseline data.

The multivariable analysis will be performed using a stepwise backward Cox proportional hazard model. Variables with a p value >0.1 will be removed from the model.

Depending on these results, we will then attempt to make a prognostic model if it makes sense to do so. For continuous variables that are being considered for inclusion in the prognostic model, the linearity assumption will be checked to see if a more complex relationship is more appropriate.

Statistical software will be used to perform the analyses. Statistical significance will be taken at the 5% level.

This plan was agreed by B Kahan, A Clive and N Maskell on February 4th 2013

NB. In order to combine the data with the Dutch and Australian cohorts, it was felt that defining survival from date of diagnosis to death would be more appropriate than trial entry to death. Analysing the data with this definition did not substantially alter the findings.

Table: Characteristics of the 221 patients in the 'UK Cohort 1'

Variable			Result	Missing Data, n
Age, Median (IQR)			74 (67-80)	0
Effusion Size, n (%)	1	Blunting of costophrenic angle	9 (4)	8
	2	More than blunting but less than 25% of hemithorax occupied by pleural fluid	22 (10)	
	3	Pleural fluid occupying 25% to 50% of hemithorax	90 (42)	
	4	Pleural fluid occupying 50% to 75% of hemithorax	61 (29)	
	5	Pleural fluid occupying more than 75% of hemithorax	31 (15)	
ECOG Performance Score, n (%)	0		24 (11)	9
	1		91 (43)	
	2		67 (32)	
	3		27 (13)	
	4		3 (1)	
Malignant cell type, n (%)	Mesothelioma		58 (26)	0
	Lung cancer		66 (30)	
	Breast/Gynaecological cancer		46 (21)	
	Other		51 (23)	
Serum Albumin (g/l), median (IQR)			33 (28-36)	4
Serum eGFR, n (%)	<15		1 (1)	0
	15-29		5 (2)	
	30-59		51 (23)	
	60-89		103 (47)	
	>90		61 (28)	
Serum NT-proBNP (pg/ml) Median (IQR)			282 (112-1001)	22
Serum NLR, median (IQR)	Absolute neutrophil count divided by the absolute lymphocyte count		5.15 (3.56-7.35)	2
Modified Glasgow Prognostic Score, n (%)	0	CRP ≤10mg/l and albumin ≥35g/l	47 (22)	11
	1	CRP >10mg/l and albumin <35g/l, or CRP ≤10mg/l and albumin < 35g/l	48 (23)	
	2	CRP >10mg/l and Albumin <35g/l	115 (55)	
Serum PLR, median (IQR)	Absolute platelet count divided by the absolute lymphocyte count		255.7 (178.3-360.0)	2
Serum CRP (mg/L), Median (IQR)			30 (11-83)	7
Pleural fluid VEGF (pg/mL), Median (IQR)			2251.5 (831.3-5722)	19
Pleural fluid LDH (IU/L), Median (IQR)			627.5 (352-1164)	7
Pleural fluid pH/glucose score, n (%)	High	Pleural fluid pH≥ 7.3 or if pH unavailable, pleural fluid glucose ≥3.3mmol/l	145 (74)	24
	Low	Pleural fluid pH <7.3 or if pH unavailable, pleural fluid glucose <3.3mmol/l	52 (26)	

Definition of abbreviations: HR = Hazard Ration; ECOG Performance Score = Eastern Cooperative Oncology Group Performance Score; NLR = Neutrophil-to-Lymphocyte Ratio; LDH = Lactate Dehydrogenase; IU/L = International units per litre; 95% CI = 95% Confidence Interval; IQR = interquartile range; VEGF = Vascular Endothelial Growth Factor; eGFR= estimated glomerular filtration rate; NT-proBNP = N-terminal prohormone of brain natriuretic peptide; CRP = C Reactive Protein; PLR = Platelet-to-Lymphocyte Ratio. * Variables with '-' were dropped from analysis due to non-significance (ie p>0.1)

Results of univariable analysis of the multiply imputed 'UK Cohort 1' dataset

Variable		Results of univariable analysis			Results of multivariable analysis		
		Hazard Ratio	95% CI	p value	Hazard Ratio	95% CI	p value
Age		1.03	1.01-1.04	0.0013	-	-	-
Effusion Size	1	1.01	0.49-2.09	0.974	1.01	0.47-2.19	0.977
	2	1.59	0.98-2.58	0.061	1.53	0.91-2.56	0.106
	3	-	-	-	-	-	-
	4	0.71	0.50-1.00	0.048	0.69	0.48-1.00	0.051
	5	0.92	0.59-1.43	0.714	0.78	0.49-1.24	0.302
ECOG Performance score	0	-	-	-	-	-	-
	1	1.31	0.78-2.18	0.304	1.42	0.84-2.39	0.192
	2	2.63	1.56-4.43	0.000	2.35	1.34-4.12	0.003
	3	3.41	1.88-6.20	0.000	2.67	1.39-5.10	0.003
	4	4.60	1.35-13.72	0.015	4.52	1.26-16.22	0.021
Malignant cell type	Mesothelioma	-	-	-	-	-	-
	Lung cancer	2.27	1.56-3.30	<0.001	1.46	0.97-2.20	0.072
	Breast/Gynecological cancer	0.87	0.57-1.34	0.529	0.79	0.50-1.25	0.314
	Other	1.32	0.88-1.96	0.178	0.96	0.62-1.48	0.858
Serum Albumin (g/l)		0.94	0.91-0.96	<0.001	-	-	-
Serum eGFR	<15	3.24	0.44-23.70	0.247	-	-	-
	15-29	1.39	0.56-3.50	0.479	-	-	-
	30-59	1.04	0.70-1.54	0.844	-	-	-
	60-89	0.82	0.58-1.15	0.248	-	-	-
	>90	-	-	-	-	-	-
Serum NT-proBNP/100 (pg/ml)***		1.01	1.01-1.01	<0.001	1.01	1.00-1.01	0.033
Serum NLR		1.08	1.05-1.12	<0.001	1.07	1.03-1.11	0.000

mGPS	0	-	-	-	-	-	-
	1	0.91	0.59-1.41	0.685	-	-	-
	2	1.66	1.15-2.38	0.006	-	-	-
Serum PLR		1.00	1.00-1.00	0.0201	-	-	-
Serum CRP (mg/L)		1.00	1.00-1.01	0.0024	-	-	-
Pleural fluid VEGF (pg/mL)		1.00	1.00-1.00	0.4130	-	-	-
Pleural fluid LDH/100 (IU/L)**		1.01	1.00-1.03	0.0368	1.02	1.00-1.03	0.010
Pleural fluid pH/glucose score		0.96	0.70-1.32	0.8000	-	-	-

Definition of abbreviations: ECOG Performance Score = Eastern Cooperative Oncology Group Performance Score; NLR = Neutrophil-to-Lymphocyte Ratio; LDH = Lactate Dehydrogenase; IU/L = International units per liter; 95% CI = 95% Confidence Interval; IQR = interquartile range; VEGF = Vascular Endothelial Growth Factor; eGFR= estimated glomerular filtration rate; NT-proBNP = N-terminal prohormone of brain natriuretic peptide; CRP = C Reactive Protein; PLR = Platelet-to-Lymphocyte Ratio;

* Variables with '-' were dropped from analysis due to non-significance

** The Hazard Ratios and 95% Confidence Intervals for pleural fluid LDH are for a 100IU/L increase in LDH.

*** The Hazard Ratios and 95% Confidence Intervals for NT-proBNP are for a 100pg/ml increase in NT-proBNP.

Table: Characteristics of the 83 patients in the 'UK Cohort 2' and the results of multivariable analysis

Variable		Result	Missing Data	Results of multivariable analysis		
				Hazard Ratio	95% CI	*p* value
ECOG Performance Score, n (%)	0	13 (16)	2	-	-	-
	1	32 (40)		3.42	0.92-12.73	0.066
	2	21 (26)		2.18	0.57-8.34	0.256
	3	15 (19)		4.65	1.20-18.11	0.027
	4	0 (0)		-	-	-
Malignant cell type, n (%)	Mesothelioma	17 (21)	1	-	-	-
	Lung cancer	33 (40)		3.09	1.10-8.69	0.033
	Breast/Gynaecological cancer	15 (18)		2.83	0.88-9.14	0.081
	Other	17 (21)		3.62	1.09-12.02	0.036
Serum NLR, median (IQR)		4.97 (3.33-7.95)	1	1.09	1.03-1.15	0.004
Pleural fluid LDH (IU/L), median (IQR)		620.5 (352.5-1153)	3	1.02*	0.99-1.04*	0.163
Definition of abbreviations: ECOG Performance Score = Eastern Cooperative Oncology Group Performance Score; NLR = Neutrophil-to-Lymphocyte Ratio; LDH = Lactate Dehydrogenase; IU/L = International units per litre; 95% CI = 95% Confidence Interval. * The Hazard Ratios and 95% Confidence Intervals for pleural fluid LDH are for a 100IU/L increase in LDH.						

Sensitivity analysis of LENT Score using the 'UK Cohort 1'

SCORE	ASSUMING BEST CASE						ASSUMING WORST CASE					
	n	HR (95% CI)	Median (IQR) survival (days)	Survived to 6 months – n (%)	Survived to 3 months-n(%)	Survived to 1 month-n(%)	n	HR (95% CI)	Median (IQR) survival (days)	Survived to 6 months, n (%)	Survived to 3 months-n(%)	Survived to 1 month-n(%)
LOW RISK (score 0-2)	49	-	319 (236-549)	42 (86)	48 (98)	49 (100)	44	-	319 (228-549)	38 (86)	43 (98)	44 (100)
MODERATE RISK (score 3-4)	140	1.61 (1.13-2.28)	125 (41-427)	62 (44)	81 (58)	112 (80)	138	1.50 (1.04-2.17)	160 (48-443)	65 (47)	85 (62)	114 (83)
HIGH RISK (score 5-7)	32	5.33 (3.27-8.67)	44 (22-77)	2 (6)	5 (16)	21 (66)	39	5.46 (3.41-8.77)	43 (22-77)	3 (8)	6 (15)	24 (62)
Harrell's C	0.6921						0.706					
AUC for 1 month mortality	0.76 (0.69-0.83)						0.77 (0.71-0.84)					
AUC for 3 months mortality	0.84 (0.79-0.89)						0.83 (0.78-0.88)					
AUC for 6 month's mortality	0.83 (0.78-0.89)						0.83 (0.78-0.89)					
Definition of abbreviations: HR= Hazard Ratio; 95% CI = 95% Confidence Interval; IQR = interquartile range; AUC = Area Under Curve NA * Unable to provide data as insufficient patients have died.												

Sensitivity analysis of LENT Score using the 'UK Cohort 2'

SCORE	ASSUMING BEST CASE						ASSUMING WORST CASE					
	n	HR (95% CI)	Median (IQR) survival (days)	Survived to 6 months – n (%)	Survived to 3 months-n(%)	Survived to 1 month-n(%)	n	HR (95% CI)	Median (IQR) survival (days)	Survived to 6 months, n (%)	Survived to 3 months-n(%)	Survived to 1 month-n(%)
LOW RISK (score 0-2)	15	-	347 (336-NA*)	13 (87)	14 (93)	15 (100)	14	-	NA* (347-NA*)	12 (86)	13 (93)	14 (100)
MODERATE RISK (score 3-4)	50	2.50 (0.97-6.45)	198 (80-NA*)	29 (58)	33 (66)	47 (94)	49	2.74 (0.96-7.81)	249 (81-NA*)	29 (59)	33 (67)	46 (94)
HIGH RISK (score 5-7)	18	8.67 (3.10-24.21)	51 (24-119)	3 (17)	6 (33)	13 (72)	20	9.97 (3.30-30.17)	51 (24-119)	4 (20)	7 (35)	15 (75)
Harrell's C	0.7385						0.7446					
AUC for 1 month mortality	0.80 (0.67-0.93)						0.77 (0.64-0.91)					
AUC for 3 months mortality	0.75 (0.64-0.85)						0.73 (0.63-0.84)					
AUC for 6 month's mortality	0.80 (0.70-0.89)						0.78 (0.68-0.88)					
Definition of abbreviations: HR= Hazard Ratio; 95% CI = 95% Confidence Interval; IQR = interquartile range; AUC = Area Under Curve NA * Unable to provide data as insufficient patients have died.												

Sensitivity analysis of LENT Score including only patients with a histocytological diagnosis of malignancy (using 'UK Cohort 1')

SCORE	n	HR (95% CI)	Median (IQR) survival (days)	Survived to 6 months – n (%)	Survived to 3 months- n(%)	Survived to 1 month- n(%)
LOW RISK (score 0-2)	44	-	319 (228-549)	38 (86)	43 (98)	44 (100)
MODERATE RISK (score 3-4)	125	1.41 (0.97-2.04)	181 (63-467)	63 (50)	82 (66)	104 (83)
HIGH RISK (score 5-7)	33	6.20 (3.76-10.23)	43 (22-60)	2 (6)	4 (12)	20 (61)
Harrell's C	0.7127					
AUC for 1 month mortality	0.79 (0.72-0.85)					
AUC for 3 months mortality	0.85 (0.80-0.90)					
AUC for 6 month's mortality	0.84 (0.79-0.89)					
Definition of abbreviations: HR= Hazard Ratio; 95% CI = 95% Confidence Interval; IQR = interquartile range; AUC = Area Under Curve						

Sensitivity analysis of LENT Score excluding patients with mesothelioma (using 'UK Cohort 1')

SCORE						
	n	HR (95% CI)	Median (IQR) survival (days)	Survived to 6 months – n (%)	Survived to 3 months- n(%)	Survived to 1 month- n(%)
LOW RISK (score 0-2)	6	-	484 (228-NA*)	6 (100)	6 (100)	6 (100)
MODERATE RISK (score 3-4)	118	2.46 (0.90-6.70)	130 (47-467)	53 (45)	70 (59)	97 (82)
HIGH RISK (score 5-7)	39	7.83 (2.73-22.44)	43 (22-77)	3 (8)	6 (15)	24 (62)
Harrell's C	0.69					
AUC for 1 month mortality	0.69 (0.60-0.78)					
AUC for 3 months mortality	0.76 (0.70-0.83)					
AUC for 6 months mortality	0.81 (0.74-0.87)					
Definition of abbreviations: HR= Hazard Ratio; 95% CI = 95% Confidence Interval; IQR = interquartile range; AUC = Area Under Curve						

Sensitivity analysis evaluating only patients with a new diagnosis of malignancy (using 'UK Cohort 1')

SCORE	n	HR (95% CI)	Median (IQR) survival (days)	Survived to 6 months – n (%)	Survived to 3 months- n(%)	Survived to 1 month- n(%)
LOW RISK (score 0-2)	42	-	319 (239-549)	36 (86)	41 (98)	42 (100)
MODERATE RISK (score 3-4)	89	1.56 (1.05-2.31)	168 (47-443)	42 (47)	54 (61)	74 (83)
HIGH RISK (score 5-7)	24	5.97 (3.45-10.33)	33 (13-57)	2 (8)	3 (13)	13 (54)
Harrell's C	0.7264					
AUC for 1 month mortality	0.83 (0.76-0.89)					
AUC for 3 months mortality	0.88 (0.82-0.93)					
AUC for 6 month's mortality	0.84 (0.78-0.90)					
Definition of abbreviations: HR= Hazard Ratio; 95% CI = 95% Confidence Interval; IQR = interquartile range; AUC = Area Under Curve						