Reviewer’s report

Title: Prognostic factors for survival in patients with malignant pleural effusions. What is considered important for treatment?

Version: 2
Date: 16 December 2014

Reviewer: Amelia Clive

Reviewer’s report:

The authors of this retrospective case series should be congratulated on their study, which evaluates a number of prognostic factors in malignant pleural effusion.

1. Major Compulsory Revisions

• The results section of the abstract should focus less on the baseline patient information and more on the important findings from the survival analyses (HR and 95% CI should be added when referring to the important findings in multivariable analysis).

• In the methods section it would be useful to know this additional information:
  # Were all patients symptomatic from their malignant effusion or were some picked up incidentally for another reason and then investigated because of a clinical suspicion of malignancy? If so, what proportion of the patients were symptomatic?
  # Did the MPE represent a new diagnosis of malignancy for all patients in the study, or did it represent progressive or recurrent disease in some cases? If so, how many patients fitted into these categories.
  # It should be clarified whether complete case data was available for all variables in all patients. If not, how was missing data dealt with.
  # Some clarification of the exclusion criteria would be beneficial. What does ‘untreated’ refer to exactly? It implies those who had thoracic surgery were still included. Does this only refer to chemotherapy?
  # Why was survival time calculated in months rather than days? It seems odd that all the survivals reported in table 2 are ‘whole’ months and I suspect that a more subtle but important potential signal could be lost by rounding survival up and down by such a large margin. I would suggest that analysing the data in terms of days from diagnosis to death would be a much more robust approach.
  # The statistical analysis section of the methods is confusing and should be made clearer. I don’t understand the sentence regarding how the discrete variables are expressed. Regarding the selection of factors for inclusion in the multivariable model, what significance level was used? In the multivariable analysis, it refers to adjusting for confounders- what exactly does this refer to? Also, what p value was used in the backwards selection model?
  # It also appears from table 2 that many of the variables were analysed as
categories using a median split to divide them. This should be stated in the methods and the reasons for analysing them in this way justified (rather than as continuous variables). I wonder whether a potentially important signal may have been missed given the fact that only 2 categories were used in the statistical analysis of the tumour type.

# The section regarding sample size should be more concise and clear as to its purpose. The results do not seem to refer to external validation with another data set, so I am unclear what the last paragraph of the statistical analysis section refers to.

• Regarding the Figures and tables:
# It would be useful to include a ‘numbers at risk’ section to the Kaplan Meier curves.
# Table 1: If there is any missing data, a column should be added to this table stating the numbers of patients with missing data for each variable. Also, I don’t understand what the neutrophil and lymphocyte rows actually refer to- this needs to be clarified.
# Table 2: Please add a column for the number of patients in each group.
# Table 3: I am not sure it is necessary to report both the Chi2 and p value. I think one would suffice and might make the table easier to read. I would also consider removing the ‘overall’ row for each variable as I think this is confusing. Also, the Hazard Ratio (95% CI) column is confusing because there appear to be figures allocated to the overall category (I think this may just be a formatting issue) but it is confusing.

• Regarding the discussion:
# The section regarding the stating of lung cancer does not seem directly relevant to the reported data.
# It would be beneficial to add a paragraph to the discussion regarding the limitations of the study, for example the retrospective nature of the series, the small number of patients in some of the categories, the lack of patients with mesothelioma in the cohort.
# A number of relevant and important, recent or large studies in this field have been omitted from the discussion and should be considered for inclusion. For example there are a number of recently published articles evaluating survival in MPE: Bielsa et al., Eur J Int Med 2008; Clive et al., Thorax 2014; Anevlas et al., Respiration 2014.
# Again, regarding the use of IPCs there is robust data from 2 RCTs which should be referred to (Davies et al., JAMA 2012; Demmy et al, JNCCN 2012). I personally don’t agree with the statement in the final paragraph of the discussion that IPC is generally indicated in patients with a survival of <6 weeks. I don’t think this is supported by the current literature.

2. Minor Essential Revisions
• In the abstract, the first sentence suggests standardization of the management of MPE. However, the data presented suggests that survival varies quite
substantially between patients and hence perhaps the term individualised management might be more appropriate.

• Please could you clarify what was meant by the final sentence of the ‘potential predictors of survival’ section regarding the censoring of subjects. This implies that all patients have died and none are lost to follow up. Is this correct?

3. Discretionary Revisions

• The details of the ECOG PS detailed in the paragraph entitled ‘potential predictors of survival’ might be better presented in a table to improve the readability of this section.

• For ease of reading, I am not sure it is necessary to explain what a hazard ratio is in the statistical analysis section. I also think the section in the discussion explaining the Cox Model could be removed.

• Rather than using the abbreviation P- in the tables, it would be easier to follow if pleural fluid was written out in full each time.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Needs some language corrections before being published

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I declare that I have no competing interests.