Reviewer’s report

Title: Validation study of a diagnostic aid to rule out pneumonia in adults with cough and fever

Version: 3 Date: 6 November 2012

Reviewer: Tom Fahey

Reviewer’s report:

Review of Held et al. Validation study of a diagnostic aid to rule out pneumonia in adults with cough and fever

This study was designed to validate a diagnostic aid to rule out pneumonia in adults with cough and fever. We have a number of issues in relation to selection of patients for the validation study, plan of analysis and conclusions drawn by the authors:

Major compulsory revisions

Methodological issues that require clarification-

• Selection bias - participants were excluded if they were prescribed antibiotics and it is unclear if participants were enrolled consecutively. The validation cohort consequently consists of a spectrum of patients at low risk of developing pneumonia and requiring antibiotics. We feel that the validation cohort contains substantial spectrum bias for these reasons, and account for the low probability of requiring antibiotics. A broader cohort of patients is needed to validate this rule.

• Verification bias, through use of a patient reported proxy measure as the reference standard and/or information from other treating physicians. With respect to the latter, it is unclear what information was retrieved from the alternative physician.

Plan of analysis-

• There are issues around the statistical approach applied in this study. Validation of a clinical prediction rules involves discrimination and calibration. Discrimination is a measure of how well a model categorises individuals into those who will or will not get the disease outcome (in this case pneumonia/requirement for antibiotic). Discrimination is measured by using the area under the receiver-operating characteristic curve (AUROC). The ROC curve is a plot of sensitivity against 1-specificity, or put another way, true positives versus false positives. Calibration is quantified by the ratio of the predicted risk to the observed risk and can be graphically displayed with a calibration curve that plots predicted versus observed outcomes. Goodness-of-fit test statistics include the C statistic the Hosmer Lemeshow test.[1]

Standardised reporting-

• Although standardised reporting guidelines for the validation of risk prediction
models are only being developed, relevant elements from other publications can be included—e.g. the STROBE guidelines and the User’s guide on methodological standards for clinical decision rules.[2][3] Taking these into consideration, the paper needs to include clear definitions of all variables—in particular the authors need to provide a definition of pneumonia, how it was measured as an outcome event and clarify and define the variables increased body temperature/fever. They need to clarify if the GPs in the follow-up study were also involved in the derivation study.

• It is unclear if there was no loss to follow-up or if the numbers of potentially eligible patients is not correctly reported. With respect to the outcome measure, it is unclear the proportion of patients in which pneumonia was defined by self-report questions and/or from another treating physician. Furthermore, there are a very low number of outcome events, which may be a direct result of selection bias. The results would benefit from the inclusion of detail around increased body temperature/fever of patients included in this study, including the mean temperature and the proportion of patients with measured temperature versus subjective self-reporting of increase in body temperature.

Additional comments:

• The terms ‘increased body temperature’ and ‘fever’ are used interchangeably throughout the paper. For example, the title indicates the aid is designed for people with fever, while the aims in the Background indicate that the aid is designed for patients with increased body temperature. This requires some clarification. A clear definition of this inclusion criterion needs to be included.

• Although the authors discuss the study limitations of selection and validation bias, they should also comment on why only one aspect of the original rule was measured and differences in populations between the two studies in terms of clinical setting, pre-test prevalence and exclusion based on treatment with antibiotics. Statistical limitations outlined above should also be considered. Further validation of the diagnostic aid is necessary before widespread application in clinical practice. The discussion and conclusion should be rewritten to reflect this. Furthermore the conclusion should acknowledge the impact of study limitations and results should be interpreted cautiously until further validation.

• There are a number of discrepancies between the derivation study and the current study, which have either not been clearly articulated or not been mentioned. In particular, the authors should explain why only one element of the original decision aid is being validated, differences in clinical settings, the exclusion of patients who receive antibiotics and differences in the reference standard measurement of the outcome event. The authors should also comment on the validity and reliability of their proxy measure of clinically relevant pneumonia relative to the reference of chest x-ray used in clinical guidelines.

• The title does not reflect the findings of the study. The conclusions in the abstract should be more balanced. The findings of this study should be interpreted in the context of the limitations of the paper. It should also reflect the fact that further validation of the aid is necessary before regular application of the
aid in clinical practice.

- Minor comments – the Background could include more information on the derivation article, including the setting and age group of patients in which the aid was derived; the Methods and Discussion sections would benefit from the inclusion of subheadings; Results, page 6, when describing second patient, suggest removing ‘her’ from sentence.

References


Level of interest: An article of limited interest

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

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I declare that I have no competing interests