Author's response to reviews

Title: Diagnostic aid to rule out pneumonia in adults with cough and feeling of fever. A validation study in the primary care setting

Authors:

Ulrike Held (ulrike.held@usz.ch)
Claudia Steurer-Stey (claudia.stey@usz.ch)
Felix Huber (felix.huber@medix.ch)
Sergio Dallafior (dallafior@pg-altstetten.ch)
Johann Steurer (johann.steurer@usz.ch)

Version: 4 Date: 22 November 2012

Author's response to reviews: see over
Point by point reply to the comments of the reviewers

<table>
<thead>
<tr>
<th>Reviewer: Tom Fahey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maior revisions</td>
</tr>
<tr>
<td>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.</td>
</tr>
<tr>
<td>We fully agree with the reviewer that we included only a selection of all patients with cough and feeling of increased body temperature in this validation study to rule out pneumonia. Only patients with CRP-levels less than 50 μg/ml, no dyspnea reported by the patient, no daily feeling of increased body temperature since the onset of cough and no treatment with antibiotics after the first consultation were included in the validation sample. To emphasize this point we added a subtitle in the methods part (In- and exclusion criteria) and rewrote the corresponding sentence of this paragraph. It reads now: Only patients in which our decision aid was negative, meaning patients with CRP levels less than 50 μg/ml (normal &lt; 10 μg/ml), no dyspnea reported by the patient, no daily subjective feeling of increased body temperature since the onset of cough, and no treatment with antibiotics after the first consultation.</td>
</tr>
<tr>
<td>After reading the thoughtful comments of the reviewers and rereading our paper we believe that a more precise formulation of the study goal at the end of the introduction section makes the paper more clear.</td>
</tr>
<tr>
<td>It reads now: Before integrating such a diagnostic aid into daily practice, the accuracy of the instrument should be tested in a new sample of patients. The aim of this study is to evaluate the accuracy of the diagnostic aid in patients with CRP levels below 50 μg/ml, no dyspnea and not daily subjective feeling of increased body temperature since onset of</td>
</tr>
</tbody>
</table>
We agree that a larger cohort is needed to further validate this rule, but we have difficulties to understand why the rule should be validated in a broader cohort (e.g., patients with higher CRP levels and or dyspnea and or daily fever). Our intention was to validate the diagnostic aid as rule-out criterion, only in those patients in which the ‘test’ result (diagnostic aid) was negative, meaning rule-out of pneumonia. Following the rule, these patients have CRP-levels less than 50 μg/ml, no dyspnea, and no daily feeling of increased body temperature since the onset of cough. In order to make inference about the proportion of false negatives in this sample, no treatment with antibiotics after the first consultation was an additional criterion the patients needed to fulfill.

We asked participating physicians to enroll patients consecutively; we describe honestly in the discussion section (limitations) that we cannot guarantee that all eligible patients were enrolled consecutively. To our experience the consecutive enrollment is almost always difficult under daily practice conditions, but the physicians told us that they did their best.

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

We agree that this sentence is misleading. In Switzerland it is possible to visit any physician different from the one attended for the first consultation. In the case it would have happened we would have contacted also the second physician. This never occurred, therefore we delete the passage … a physician different from the one at the first visit… from the sentence to make it clearer.

We changed the sentence and it reads now: When a patient had been hospitalized since inclusion in the study, we contacted hospitals to obtain the information whether pneumonia or another illness was the reason for the hospitalization.

There are issues around the statistical approach applied in this study. Validation of a clinical prediction rules involves discrimination and

Please let us clarify this point: the design of this validation study was to validate a rule-out criterion. The criterion was derived from a classification tree, and not from a logistic regression model and thus does not intend to predict an individual patient’s probability for pneumonia. For that reason, we only
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.

Although standardised reporting guidelines for the validation of risk prediction models are only 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. Taking these into consideration, the paper needs to include clear definitions of all variables included patients for which our test criterion was negative meaning rule out of pneumonia. If you think in terms of a 2x2 table, we did not collect patients in the row were the test criterion equals one (T=1), but only in the row where it equals zero (T=0). In this population of included patients, the assessment of AUC, calibration and Hosmer-Lemeshow test statistic, but also the calculation of sensitivity and specificity is not possible. We were mainly interested in making inference about patients in which our rule was negative, and estimating the proportion of “false negatives” among them.

Many thanks for these helpful recommendations:

Pneumonia: In the derivation paper pneumonia was defined “as a set of symptoms and signs consistent with an acute lower respiratory tract infection associated with radiographic shadowing for which there is no other explanation.”

This paper focuses on ruling-out pneumonia. As we were not able to perform x-rays in all these patients (all of them with a low probability of pneumonia) we had to apply a proxy measure. We added the description of this proxy measure in the Method section (paragraph before statistical analysis).
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.

When patients at the follow-up consultation (by phone or in the office) reported that symptoms improved and no antibiotics had been prescribed to them since the first consultation we assumed that a clinically relevant pneumonia could be excluded with a high degree of certainty. In the discussion section we describe the potential limitations of this proxy measure.

Increased body temperature: body temperature was not measured in these patients in the GP’s practice and most of the patients did not measure their temperature at home; the variable is defined as the subjective feeling of increased body temperature. We changed in the manuscript to *subjective feeling of increased body temperature*.

Gp’s participating in the derivation study were not involved in the validation 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 | All 110 enrolled patients could be contacted by phone. Eligible were 115 patients but 5 patients were not included because they received a prescription for antibiotics at their first consultation.

To emphasize the fact, that there was no loss to follow-up we changed the sentence in the result section (third paragraph), and it reads now; *All 110 patients enrolled in the study were contacted after a median of 13 days.*

In no patient a diagnosis of pneumonia was established. We agree as written above that we might have missed patients with pneumonia but the probability of having missed a “clinically” relevant pneumonia is very low.

We agree that the low number of outcome events is the result of the selection of patients. We have no results of temperature measurements before the first consultation from these patients. We also refrained from measuring body temperature at the time of consultation, because many of the patients take some antipyretic substance and therefore temperature at the time of consultation is not easy to interpret. |
patients with measured temperature versus subjective self-reporting of increase in body temperature.

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.

We acknowledge that the usage of these two terms is confusing. We changed in the whole manuscript to the term “subjective feeling of increased body temperature”. The title was changed accordingly.

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

By describing the aim of the study inclusion criteria more precisely we hope and believe that these concerns could be resolved. Our aim was to evaluate the accuracy of the diagnostic aid in a specified subset of patients with cough and subjective feeling of increased body temperature. We agree that the pre-test prevalence is very low in our population. We do not fully understand what the reviewers mean by ‘only one aspect of the original rule was measured’. The intention of this study was to evaluate the decision aid as rule-out criterion only in those patients in which the rule was negative (indicating no pneumonia).

We changed the abstract, and the Conclusions accordingly and it reads now.

This diagnostic aid is helpful to rule out pneumonia in patients from a primary care setting. After further validation application of this aid in daily practice may help to reduce the
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.

<table>
<thead>
<tr>
<th>prescription rate of unnecessary antibiotics in patients with respiratory tract infections.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>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.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>To clarify: The aim of this study was to validate the rule out aspect of the decision aid, and therefore a subset of all patients with cough and increased body temperature was included. We made inference about the proportion of patients with pneumonia among those with negative test criterion. It was not our intention to validate the regression model for estimating the probability of pneumonia in all patients with cough and subjective feeling of increased body temperature. From the first study (derivation study) in addition to the development of a regression model to estimate the probability of pneumonia we derived a diagnostic aid based on a classification tree to rule out pneumonia in patients with cough and the subjective feeling of increased body temperature. We assume that the diagnostic aid is more useful in daily practice than the calculation of probabilities and therefore focused on the validation of the rule-out instrument.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>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</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>We changed the title Diagnostic aid to rule out pneumonia in adults with cough and feeling of fever. A validation study in the primary care setting … and adapted the conclusions in the abstract and in the main manuscript. We fully agree that further validation is necessary before recommending the broad application of the diagnostic aid.</th>
</tr>
</thead>
</table>
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.

We added more information on the derivation study in the introduction of the paper and it reads now. A
total of 621 patients, almost all of them attending a primary care physician, were included in the
derivation study. The mean age of those patients was 48 years and half of them were males. The
derivation study showed that three pieces of information are necessary to rule out pneumonia in these
patients;

We added subheadings in the method and discussion sections: Methods; Recruitment of physicians,
In-and exclusion criteria,
Data gathering. In the Discussion section; main results, limitations of the study, clinical implications,
implications for further research.
We removed ‘her’

Reviewer: Carl Llor

The definition of the health problem is not
clearly defined. You recruited patients with
cough but not as the main symptom. We cannot
therefore affirm that the patients included in
your sample were diagnosed with lower
respiratory tract infections, since other
infections such as common cold and sore throat

Many thanks for this important consideration. We agree with the objection of the reviewer and we
specified the health problem and added to the corresponding sentences that cough, like in the
derivation study, was the main symptom. We agree that patients with common cold and sore throat
were probably included, but in all of them cough was the main symptom that prompted the
consultation.

When esteemed reviewers and editors agree we would like to refrain from a discussion about the
potential illnesses of the patients in the study. We have not collected enough information to say and
might have been included in your study. You report that the median duration of cough in your sample was seven days. This finding could also explain that infections other than lower respiratory tract infections were also taken into account. Please discuss this more in depth.

What does ‘subjective or measured increased body temperature’ mean? Please define. In addition, I would like you to discuss whether patients were instructed to measure the temperature daily or not in more detail.

Patients were supposed to be contacted by phone only once, fourteen days after having been included but I am not sure about this. Could you also better state the follow-up of patients in this study?

You excluded patients with known chronic lung diseases except for chronic bronchitis. Were patients with spirometrically-based COPD included or excluded?

You should also explain more clearly how many chest X-rays were performed in this sample.

write more about the potential illnesses of the included patients after pneumonia was ruled out with a high degree of certainty.

We fully agree that these terms (subjective or measured increased body temperature) may be misleading. The criterion was a subjective feeling of increased body temperature only. Body temperature was not measured in these patients during consultation. We changed all passages in the manuscript to “subjective feeling of increased body temperature.”

We added a sentence in the Method section to clarify. *Body temperature was not measured during consultation and patients were not instructed to measure body temperature at home.*

We reformulated the corresponding sentence and it reads now; *At a minimum of one week after the first consultation (inclusion of patients), physicians or physician assistants contacted the patients for follow-up only once by phone or during a planned office consultation.*

Patients with and without spirometrically verified COPD were included. We have no information in how many patients COPD has been verified by spirometry. Only six patients included in the study had chronic cough (more than 4 weeks), we added this information in table 1.

Decisions on diagnostic procedures were left to the discretion of the treating physician as is written in the Method section. A chest x-ray was performed in two patients at baseline.
<table>
<thead>
<tr>
<th>Suggested Change</th>
<th>Author's Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pneumonia can only be ruled out when the radiographic study is negative. The conclusion of this study may be misleading. Even though you discuss the fact that chest x-ray was not systematically ordered as the main limitation of this study, you should explain how many plain x-rays were performed.</td>
<td>After baseline visit in none of the patients, even in those three patients, which received a prescription for antibiotics after the initial consultation a chest x-ray was performed. To make this clearer we added the following sentence in the third paragraph of the Result section. <em>In none of the patients a chest x-ray was performed after the first visit.</em></td>
</tr>
<tr>
<td>Neither is the selection of the sample clear. Could you provide more information in this paper?</td>
<td>We agree that the selection of the sample was not described clearly: We now added in the Method section <em>Only patients in which our decision aid was negative, meaning patients with CRP levels less than 50 μg/ml (normal &lt; 10 μg/ml), no dyspnea reported by the patient, no daily subjective feeling of increased body temperature since the onset of cough...</em></td>
</tr>
<tr>
<td>I highly recommend the inclusion of a table with the main characteristics of the patients included in the study.</td>
<td>We agree and inserted table 1 with the main characteristics.</td>
</tr>
<tr>
<td>You say that antibiotics are only recommended for patients with moderate-severe chronic bronchitis. However a recent paper reports the efficacy of antibiotic therapy also for exacerbations of mild to moderate COPD (Am J Respir Crit Care Med). Please change</td>
<td>Many thanks for this important reference. We added the reference in the introduction section.</td>
</tr>
<tr>
<td>How was dyspnoea defined? Did you use a</td>
<td>We added in the Method part, <em>patient reported dyspnea</em>. We asked patients if they experienced</td>
</tr>
<tr>
<td>dyspnoea score or was it measured only by asking the patient about breathlessness?</td>
<td>breathlessness in the time since onset of cough and used no dyspnea score.</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>I highly encourage the use of numbers – not only percentages - when you describe the main results (last paragraph of page 6).</td>
<td>Fully agree. Now numbers and percentages describe the main results now.</td>
</tr>
</tbody>
</table>