Author's response to reviews

Title: Pretest expectations strongly influence interpretation of abnormal laboratory results and further management

Authors:

Paul H.H. Houben (paul.houben@hag.unimaas.nl)
Trudy van der Weijden (trudy.vanderweijden@hag.unimaas.nl)
Bjorn Winkens (bjorn.winkens@stat.unimaas.nl)
Ron A.G. Winkens (ron.winkens@hag.unimaas.nl)
Richard P.T.M. Grol (r.grol@iq.umcn.nl)

Version: 2 Date: 27 October 2009

Author's response to reviews: see over
Dear editor,

Thank you very much for inviting us to resubmit a revised manuscript, addressing the criticisms outlined in the reviewers’ reports. We have marked the changes that we have made to the manuscript and have attached our responses to each of the points made by the reviewers.

On behalf of all authors,

Yours sincerely,

Paul Houben
Points made by reviewer William Hamilton

Comment 1: The paper addresses an important question: do doctors in primary care act in a Bayesian fashion? This is not how the authors have structured their paper, and the way they have done it is fine. However, as the results suggest that GPs do act in a Bayesian fashion even if they do not make formal calculations - I would have expected more about Bayes' theorem in the introduction.

Answer 1: We have revised the introduction and now give more explicit attention to Bayes’ theorem. See also our answers to Peter Wyer’s comments (comment 5).

Comment 2: The only major point I wish to make is that the analysis is arguably too simple. Some of the tests will have yielded a continuous variable (haemoglobin for example). It is very likely that the distribution of haemoglobin results is different from the two groups (low pre-test and high pre-test). It is likely that more of the values in the low pre-test are very close to the reference threshold supplied by the laboratory. Thus if GPs were arbitrarily assigning a different threshold for abnormality, then more of the low pre-test group would be normal by the GPs personal threshold. Is it possible to test this point with either haemoglobin, glucose, or perhaps creatinine?

Answer 2: The distribution of haemoglobin results differs somewhat between the group of patients with a high pretest probability and the group with a low pretest probability. For instance, the mean haemoglobin concentrations for the two groups of anaemic women were 6.8 mmol/l (SD 0.6) and 7.1 mmol/l (SD 0.3), respectively. As such differences in test results may have influenced the results of our analysis, we now discuss this in the discussion section on page [11].

Points made by reviewer Peter Wyer

Comment 1: Since practitioner behaviour is the subject of the study, the unit of action is the individual practitioner and the effective “N” of the study is therefore 87, not 1253. This could reduce the power of the study, depending upon how the analysis was performed.
Answer 1: Since this study focused on the diagnostic and management decisions made for individual patients, we performed the analyses at patient level. We have not analysed a possible clustering of specific interpretations and behaviour at physician level, as the aim of the study was not to explore the differences in interpretation between physicians, but to explore how pretest expectations influence the interpretation of results at the level of individual patients. We recruited a large group of practitioners (87) to ensure the external generalisability of our findings. We have added a comment on this decision to the discussion section, page 12.

Comment 2: Beginning with the abstract, the true focus of the study is unclear in the presentation and does not become clear until the reader is substantially into the methods section.

Answer 2: We have rewritten the background section of the abstract to better describe the true focus of the study.

Comment 3: The objective stated here (in the abstract) is also at odds with the objective as stated at the very bottom of P 3, which appears to reflect the actual objective of the study. The abstract states the objective to be “examine the influence of pretest probability estimates and reasons for ordering tests on test result interpretation and further management”. This is very misleading and it is not until much later in the article that the reader discovers that patients whose test were ordered for reasons other than testing a specific diagnostic hypothesis were EXCLUDED from the study!

Answer 3: We have rewritten the objective and background section of the abstract, giving special attention to clarifying the objective.

Comment 4: The Methods segment of the abstract is much too sparse and does not allow the reader to determine what the study was really about. Much of the results section appears to emphasize results tangential to the true objective of the study being reported. Likewise with the conclusion.

Answer 4: We regret that the abstract was not clear according to the reviewer. We are not sure what exactly is the problem, as no specific examples are mentioned. Nevertheless, we
have tried to improve its clarity by critically rewriting the methods, results and discussion sections of the abstract.

**Comment 5:** The introduction similarly wanders through multiple aspects of testing in primary care settings and only at the very end settles down to a more nearly clear statement of the “true” objectives of the study. A much more focused intro would confine itself to tests ordered for the purpose of either confirming or refuting specific diagnostic hypotheses and would address the difference between a dichotomized diagnostic result (positive or negative) and a continuous variable with respect to the approach to interpretations and actions by non-academic clinicians.

**Answer 5:** We have revised and reduced the introduction section of the manuscript. We now focus more on the objective of the study, discussing pretest probability and the diagnostic reasons for ordering tests that we included (i.e. excluding and confirming diagnosis as well as patient reassurance), see page [4]. We also comment on Bayes’ theorem as suggested by the other reviewer (comment 1). We have tried to avoid too much discussion of the many aspects of testing. We prefer not to address differences between dichotomized and continuous results in the introduction section, as the study focused on pretest expectations influencing posttest variables, not on specific types of test results that may influence posttest variables. However, as this is an important aspect of testing, we now comment on it in the discussion section on page 12. See also our response to comment 9.

**Comment 6:** The Methods elaboration settles down to a fairly focused and consistent description of a study restricted to the issue of testing in relationship to specific disease hypotheses. Because of the unclarity of the earlier portion of the manuscript, the statement at the bottom of P 5, that patients for whom tests were ordered for reasons other than hypothesis testing were excluded, comes as a surprise to the reader and might even be entirely overlooked.

**Answer 6:** Because the statement on P 5 might come as a surprise, the introduction now states more clearly that our study concerned tests ordered for diagnostic reasons, such as confirming or excluding a diagnostic hypothesis.
Comment 7: In general, the methods would be illuminated substantially if the actual survey form administered to the practitioner subjects were provided with the report. As examples of this: P. 5, Par 2: In connection with the query “Do you suspect a disease?” were the subjects asked to name the particular disease they suspected. If not, the classification of the test ordering behaviour might have been subject to ambiguity. P. 5, Par 4: in line with the preceding query, was there any control on whether the disease being suspected at the point of the post test survey was the same disease being suspected at the point of the pre test survey?

Answer 7: We have added a translated version of the actual survey forms. At the question “Do you suspect that the patient has a disease?” – i.e. the pretest estimate of disease probability – the physicians also named the particular disease(s) they expected on the pretest survey form. The posttest survey also asked for both the physicians’ suspicion of disease – i.e. the posttest estimate of disease probability – and the particular disease(s) they suspected at that moment. However, we did not analyse whether the physicians mentioned the same disease(s) in the pretest survey as in the posttest survey, as the study did not focus on the specific diagnoses generated by the physicians, but on the influence of pretest expectations in general on test results interpretation, estimates of posttest disease probability and further management.

Comment 8: P. 5, Par. 5: This aspect of your inquiry appears plausible and potentially useful but could be explained more clearly here and elsewhere. I.e. you were apparently trying to classify and tabulate the ‘psychologies of anticipated action/management’ associated with practitioners pre test probability estimates when specific diseases were in question. Implicitly, these pretest estimates therefore were tied to anticipated results. E.g. “I think the likelihood is very low and expect that the test result will be normal or near normal. I therefore perceive myself to be ordering the test for purpose of ‘reassurance’.” A potentially hidden variable here is the extent to which patient anxiety/apprehension played an independent role in motivating the ordering of the test. Was there any attempt to assess this aspect of the process?

Answer 8: We have not assessed the role of patient anxiety in the motivation for ordering tests. This would have been very interesting, but would imply even more complex data collection. Our study was about physician-related influences on the interpretation of diagnostic test results.
Comment 9: P. 9-10. Your discussion of the implications of your findings might be enhanced if you supplemented your reference to Bayesian theory with acknowledgment of the less abstract concept of likelihood ratios associated with continuous diagnostic test results as direct measures of impact of test results on pre-test probability. As a test result moves continuously towards more extreme values, the likelihood ratio also continuously increases. Hence, the clinicians’ instincts and trends towards interpretation and action that you observed are concretely in line with this simple numerical principle and with the previous literature on this subject. Furthermore, use of this concept does not require calculations but only an understanding of the principle and simple rules of thumb regarding interpretation of likelihood ratios within certain ranges.

Answer 9: We agree that likelihood ratios are an important concept for test interpretation (1). In the discussion section, we therefore now supplement our reference to Bayesian theory with the concept of likelihood ratios. We discuss that it is believed that likelihood ratios may help physicians to interpret results, although there is some evidence that they may not always lead to better estimations of posttest probability (2, 3), see pages 12 and 13.

Points made by the associate editor

Comment 1: Since doctors interpretation of test results will depend on how far the result deviates from the normal range, the authors should comment on the impact of the magnitude of abnormalities on doctors’ interpretations.

Answer 1: In the discussion section, pages 12 and 13, we comment on the impact of the magnitude of abnormalities on doctors’ interpretations.

Comment 2: Please clarify ethical approval, including the name of the body which gave approval, with a reference number where appropriate.

Answer 2: The Medical Ethics Committee of the Maastricht University Medical Centre has given ethical approval with reference number MEC 03-195-1. We clarify this in the manuscript text on page 6.

Comment 3: Please acknowledge anyone who contributed towards the study.
Answer 3: We now acknowledge the physicians and clinical laboratories that participated in the study on page 13.

Comment 4: We recommend that you copyedit the paper to improve the style of written English.

Answer 4: We have had the paper copy-edited.

References: