Author's response to reviews

Title: Does clinical examination aid in the diagnosis of urinary tract infections in women? A systematic review

Authors:

    David Medina-Bombardó (dmedina@ibsalut.caib.es)
    Antoni Jover-Palmer (ajover@ibsalut.caib.es)

Version: 4 Date: 9 March 2011

Author's response to reviews: see over
Palma de Mallorca, 8th March 2011

Dear Editor:

I am pleased to answer as requested in your email of data February the 14t of February 2011, to the reviewers and editor requirements to the manuscript MS 12892326124440974 “Does clinical examination aid in the diagnosis of urinary tract infections in women? A systematic review.” of David Medina-Bombardó i Antoni Jover-Palmer.

Editor Requests:
- We included authors emails in the cover letter
- We included the context information within the abstract section.
- We have provided email information of both authors in the title page
- We have included the competing interest section as well acknowledgements and authors contribution’s.

Report Reviewer: Guido Schmiemann (revision date 6 December 2010)

Method: (Major revision)
- With Medline and Embase the authors concentrated on the most relevant databases. Although systematic reviews and meta analysis were excluded from the search the Cochrane Database should be searched as well – at least to find studies within references of systematic reviews that were missed with the search used.

We mainly looked for original papers. When we obtained a systematic review or meta-analysis, we systematically review all references in order to select articles as a starting point, which identified other studies. In spite we did not search in the Cochrane Database we were able to identify the only one systematic review.

– The search is two years old (December 2008) – a more recent search / an update should be conducted as relevant articles have been published since (especially the work of Paul Little and colleagues).

We find the work of Little very interesting and we realize it would be very adequate to include this paper on our review. The thing is that we have done our systematic review during 2009, with originals edited till December 2008. The manuscript was send to the BMC in August 2010 and we have had no time to include it.

Exclusion criteria: (Major revisions)
- According to Figure 1 bacteriological findings lead to exclusion of some studies (depending on bacterial count) there is no explanation in the text. Why should a study with bacterial count >10(3) be excluded?

Exclusion criteria used for this systematic review did not include any cut off point for UTI definition. A “posteriori” in most of the identified studies used as cut off >10(5). Both authors agree to exclude those studies defining UTI as bacterial count <10(5). This decision was taken in order to avoid heterogeneity of study population and disease definition as well as statistical power. We added a paragraph in the exclusion criteria section.
- Part of the search strategy was, that “the decisions made by the reviewers” were discussed. Does this mean that all abstracts/articles were independently assessed by two (?) reviewers? If so, this should be explicitly stated as a demonstration of a thorough process. The manual search for papers is not described very clearly – were the references of all included articles searched for further articles?

Effectively, each author independently reviewed selected papers and abstracts and afterwards each selection was discussed by the two authors. Definitive inclusion was done when consensus was attained. See changes in the manuscript. Also we tried to better explain “manual” searches.

Running title: (Discretionary revisions)
- I suggest changing the title since the review focuses on history taking rather than clinical examination to diagnose a UTI.

As far as we are concerned, clinical examination is a concept that includes patient assessment through history taking as well as physical examination as proposed by David Sackett Brian Haynes and Peter Turgwell (1985) in Clinical Epidemiology: A Basic Science for Clinical Medicine. We thought this concept was the more adequate for our objectives.

Method: (Major revisions)
- Neither method nor the results of a quality assessment are described in detail. What gold standard was used (did all patients in both groups receive a urine culture, what cut-off was used for diagnosis, some articles were excluded because of a Gold standard of >10(2) and 10(3) cfu/ml). Are there any differences in clinical presentation depending on the cut off values of the gold standard?

All studies included had to explicitly defined a gold standard consisting on culture results. The considered gold standard for our review was positive culture defined as bacterial count cut off in CFU/mL units. Gold standard for UTI confirmation have changed over time. The gold standard most frequently used by most of the studies was considered.

It can be expected that with a decrease in the cut off point changes in the results could appear. As we mentioned above, we gave priority on the studies homogeneity and statistical power.

- Was the consecutive recruitment of patients an explicit inclusion criterion?

Yes. As we mentioned in the methods section, one of the exclusion criteria was non consecutive patients. As recommended by Jaeschke R et al. JAMA 1994;271:389-1, patients included in studies on diagnostic tests, should be representative of the spectrum of patients in which tests will be applied. Non consecutive patients result in non representative samples and consequently in biased findings.

Analysis (Minor revision)
- “The likelihood ratio (LR) describes how many times more likely a person with disease will receive a particular test compared with a person without disease.” I disagree with this definition – the LR is a characteristic of the test under examination – the result is the probability of this patient regarding the diagnosis – not that this patient will receive a test.

We agree with your comment and we have changed the definition in the manuscript as:
“ The likelihood ratio (LR) is the ratio of two probabilities, namely the probability that a specific test result is obtained in patients with the disease divided by the probability of a test result in patients without the disease.”
Analysis (Major revisions)

- When heterogeneity was found, the threshold effect was analyzed using the Moses-Shapiro-Littenberg method. I cannot judge the appropriateness of this test.

This method is one of the most used to assess the threshold effect, also it has been used in the analysis of the DOR and ROC curves. We have not included in the manuscript because of space problems and to facilitate clinicians a much more easier reading. If you consider more convenient to include it we will be very pleased.

- Why did the authors choose a fixed effects method rather than a random effect recommended for assessing heterogeneity?

In fact we used both methods depending of the heterogeneity identified. In all but five findings, the heterogeneity is explained by the threshold effect. In five findings threshold effect can not be evaluated and heterogeneity has been assessed through chi-squared, inconsistency index and DOR. And in two findings, we estimated the LR by the random effects model and else with the fixed effects model.

We added a new table (table 3) were appear all this results.

- The authors state that “heterogeneity was assessed via Chi-square tests and I² inconsistency tests” – what are the results of the tests? Did the results of I² support the use of a fixed effects model?

See new table 3 results of heterogeneity and inconsistency results is better explained and complete the results explained in results (paragraph 7). Minor changes in the paragraph has been introduced.

- According to the attached file a study by Verest et al was included; on what grounds was this study included? It is the only study without information on a clinical sign that has been evaluated.

Effectively, this is the only study where only assesses results of urine dipstick test. The reason of inclusion was that it accomplished all inclusion and non exclusion criteria.

Discussion: (Minor revision)

- “Randomized samples would be ideal; however, there have been [no ??] published studies of diagnostic tests based on randomized samples.” Here a word seems to be missing.

We have added the missing word

Conclusion: (Major revision)

- The review suggests that clinical findings are less important than previously thought- the conclusion seems to be very farfetched. The consequence cannot be to rely a therapeutic decision on results of dipstick results alone. This would lead to overtreatment of asymptomatic bacteriuria – symptoms and the clinical presentation are mandatory for the diagnosis of a urinary tract infection otherwise we would treat “pathologic dipsticks” instead of patients.

As clinicians we completely agree with you. We have done some changes in order to better explain the meaning of the results.

- When judging the poor LR of the different findings one has to keep in mind the high prevalence
of urinary tract infection in the included studies. This high rate is at least partly the result of a selection by the treating physicians (on the ground of patients history and/or clinical presentation).

In our opinion the poor LR could be due to the difficulty of establish the same measurement standards: that is the observers agreement in clinical examination. The variables included in the different studies are based in values of high subjectivity of the clinicians involved in the measurements. In any of the studies reviewed reproducibility has been assessed.

**Report reviewer: Ellen Stobberingh** (revision date 19 December 2010)

**Major comments:**

− The articles selected for the analysis included all papers dealing with female patients presenting with symptoms of a UTI to their GP. No differentiation was made between complicated/uncomplicated, pre-versus postmenopausal, low or high UTI. What was the reason not to make a distinction. Is it to be expected that the results will be different when the different groups will be analysed separately. Please comment.

We did not included, in the literature research, an specific research for subgroups of patients. Even though, we tried to classify results by some of the subgroups. Finally it was not possible due to inconsistent or insufficient information in reviewed papers. Consequently, to stratify by groups or UTI characteristics was not possible. We agree with reviewer comments that results could be different depending of those subgroups.

− The articles were selected based on a previously defined selection protocol: 
  i) the protocol was unfortunately not provided,

Defined protocol mainly refers to search strategies, steps for paper selection and studies inclusion and exclusion criteria. This terms are specified in the methods section. Perhaps we have to consider to exclude the term ‘defined selection protocol’.

  ii) for the selection of the articles a stepwise approach was followed. Step 5 included the judgement of an external expert whereas on the other hand articles were excluded when there was no agreement of 2 or more reviewers. Could you explain what the role of the external expert is, and do I understand it correctly that if one reviewer agreed, the article will be included?

In case of discrepancies between the reviewers, an external expert reviewed the study and the two concordant decisions were those considered. Sentence has been modified in order to better explain the process.

**Results:**

− In the last paragraph of the results the authors mentioned that data could not be analysed as only 2 studies assessed these variables. As these variables are relevant clinical findings, what is the effect of omitting them from the analysis? Are the results obtained then still valid?

We thing that it could be a misunderstanding of the sentence. In this sentence we were referring to the fact that we were unable to analyse the threshold effect. The consequences are a higher possibility of unstable results if more studies on the same findings were added. We specify in the paragraph the analysis of threshold effects.
Minor comments:

Inclusion criteria:

Please explain:

i) "agar plate culture method" or give a reference,

All studies mentioned this culture method as gold standard but now one gave more details. We added ‘usual’.

ii) how long is a "sufficient" follow-up method,

We specified better in the text.

iii) what is a "poorly defined population", which criteria were used,

It was considered ‘Poorly defined population’ when in the methods section of each paper, there was not information about context, gender, age or when inclusion criteria of study subjects were not mentioned. Find changes in the manuscript.

iv) which definition is used for "not of recent onset"

We have specified the definition of “not of recent onset” in the manuscript

v) what is the reason to exclude studies with non-systematic assessment of clinical findings (which criteria were used), non-consecutive or non-randomized recruitment of patients.

We considered non-systematic assessment of clinical findings when authors did not describe specifications of materials and methods involved, including how and when measurements were taken. Added to manuscript.

Furthermore, as recommended by Jaeschke R et al. JAMA 1994; 271:389-1, patients included in studies on diagnostic tests, should be representative of the spectrum of patients in which tests will be applied. Non consecutive patients result in non representative samples and consequently in biased findings.

The first and third paragraph of the results are part of the Materials and Methods section.

In the first paragraph we detail the results of the bibliography search and in the third the ones of the meta analysis.

Reviewer report: Malinee Laopaiboon (revision date 12 February 2011)

1) This is a systematic review paper, quality assessment of the included studies is an important component of the methods. The authors did not have this in the method part. The quality assessment results for each included study will also be needed in the result part.

We really did not rank the quality of the studies. Moreover with inclusion and exclusion criteria we have taken into account quality criteria for papers inclusion.

2) There are many clinical outcomes presented in the results. The authors should present clearly
number of included studies and subjects for individual outcomes. This information is not available in the results.

Find included in table 3.

3) The authors should provide setting and country of each included study in Table 2. Column 4 of the table should name as 'tests and reference'.

Setting and country of each study has been included in table 2. All studies have been carried out in primary care. We handed variable measures by tests and reference.

4) The authors should label figure 2, 3 and 4 what the two colors on each bar are.

Done.

5) The authors should not present results in the discussion part. Please consider paragraph 6 of the discussion. It should be in the results.

As the reviewer suggested, we moved paragraph 6 to the results section.

We tried to cover all reviewers comments and include changes in the manuscript as well as come changes in format and references.

We are looking forward this new version would satisfy the reviewers and permit acceptance for publication.

Yours sincerely,

David Medina-Bombardó

Antoni Jover-Palmer