Author’s response to reviews

Title: Evaluation of dipstick analysis among elderly residents to detect bacteriuria: a cross-sectional study in 32 nursing homes

Authors:

Pär-Daniel Sundvall (par-daniel.sundvall@vgregion.se)
Ronny K Gunnarsson (ronny.gunnarsson@vgregion.se)

Version: 2 Date: 8 April 2009

Author’s response to reviews: see over
We appreciate the time and effort the reviewers spent on studying and commenting our manuscript. Below are the reviewers' comments (reviewer #1 and #2) and our corresponding responses.

#1: Abstract: The abstract covers the article, only the last sentence in my opinion is a statement that cannot be built upon the results of the study, but is a physicians’ opinion.

Our response: We agree with the reviewer that the last sentence was an interpretation rather than a direct finding. We have deleted the last sentence.

#1: Background: I have only one comment on this section. The authors state that 'it is unclear what clinical features or events are relevant in bacteriuria' implying that the study of Juthani-Mehta focusses on a group of patients that is not relevant. I disagree with the authors on this point: dipstick analysis is done in clinical practice after features/events are mentioned to the physician: physicians will not do dipstick analysis routinely (for example every six months), but do this because signs and symptoms like the ones used by Juthani-Mehta, are present.

Our response: This is a very interesting and important viewpoint from the referee. The ultimate goal of a study should be to estimate the probability that a urinary bacteriological finding represents the aetiology of a specific symptom. This would require a gold standard that in patients with symptoms, for example changed mental status and bacteriuria, can differentiate between those with urinary tract infection causing changed mental status and those with a bacteriuria unrelated to the change in mental status. Unfortunately, there is no such gold standard. We usually try to solve this by assuming that bacteriuria in symptomatic patients is related to the symptoms. However, the scientific evidence for this assumption is weak. Specific symptoms usually associated with urinary tract infection are dysuria, change in voiding pattern or fever. However, these symptoms were not common in Juthani-Mehta’s study (7%, 6% and 12%). Most patients in Juthani-Mehta’s study had more unspecific symptoms such as change in mental status (40%), change in behaviour (20%), change in character of urine (17%), evaluation for other infections (7%) and family or patient requests (7%). In clinical practice, as stated by the referee, these unspecific symptoms often result in dipstick analysis. However, the scientific evidence for these unspecific symptoms actually being caused by a urinary tract infection is weak. One supporting finding is the prevalence of bacteriuria among asymptomatic nursing home residents is approximately the same as the prevalence in Juthani-Mehta’s and our study. Thus, our intention is not that the unspecific symptoms used by Juthani-Mehta are irrelevant but rather that we presently do not know if the unspecific symptoms are relevant or not.

Since PPV and NPV depend on prevalence of bacteriuria there should be no major differences between evaluating symptomatic or asymptomatic individuals. The focus in our study was to further clarify PPV and NPV of dipstick analysis to predict bacteriuria and not to evaluate the relevance of different symptoms.

We agree, however, with the referee that routine analysis every six months is an irrelevant procedure. We have deleted most of this text in the introduction. Second
paragraph in subsection Methodological aspects (in Discussion) has been rewritten to comment on this.

#1: Methods: the choices made in selection of patients and in use of gold standard, are clearly explained.

Our response: We appreciate this comment.

#1: Discussion: Methodological aspects: The authors state 'In this study 32% of urine cultures showed growth of potentially pathogenic bacteria. If only individuals with symptoms indicating a possible urinary tract infection had been included the results might have differed slightly. However, this study focused on evaluating dipstick ability to predict bacteriuria, not urinary tract infection. Furthermore, the major problem with selecting symptomatic individuals was that different opinions existed regarding symptoms among elderly and their correlation to urinary tract infection [1, 24]. Due to these problems individuals in this study were included regardless of symptoms.' I disagree on this point with the authors: I think it would especially be worthwhile to evaluate the NPV/PPV of the dipstick in patients with symptoms of a possible UTI, because this resembles common clinical practice: dipstick testing is used as first diagnostic step after symptoms/signs are presented by the patient to the physician. The symptoms/signs used by Juthani-Mehta are 'consensus' symptoms/signs, that are used by many clinicians.

Our response: Prevalence of bacteriuria among asymptomatic residents in nursing homes for the elderly is high and similar to the prevalence found by Juthani-Mehta et al (40%) and this study (32%). Since PPV and NPV for dipstick analysis depend on prevalence of bacteriuria there should be no major differences between evaluating dipstick analysis on symptomatic or asymptomatic individuals.

#1: The authors focussed 'on evaluating dipstick ability to predict bacteriuria, not UTI'. However they used not one gold standard, but different gold standards depending on presence of symptoms/signs like the ones used by Juthani-Mehta: this is strange in the light of their chosen focus. In my opinion, analysing the data in two ways, could provide more valuable information:
1. Analysis of all dipsticks using a culture with growth >=100.000 CFU/ml as gold standard. This analysis could provide the answer to their main focus: ability of dipstick analysis to predict bacteriuria in nursing home residents. (A problem with this analysis probably will be that the cultures of the residents with symptoms of possible UTI have been judged only for >= 1.000 (male) and >=10.000 (female)).
2. Analysis of dipstick of residents with symptoms/signs of UTI, using a culture with growth >= 1.000 CFU/ml (male) and >=10.000 (female) as a gold standard. This analysis could provide an answer to another important topic: value of dipstick analysis for predicting in bacteriuria in nursing home residents with possible UTI.

Our response: We apologise for being unclear when explaining the different cut-off points for a culture to be considered positive. The only symptoms/signs that, for some of the bacteria, lower the cut-off point are fever, frequency, urgency, dysuria, a
positive nitrite dipstick or a leukocyte esterase dipstick >1. Unspecific symptoms such as change in mental status, change in behaviour, change in character of urine, evaluation for other infections and family or patient requests etc were not symptoms that changed the cut-off point. We have now clarified this in the methods section.

We found it valuable to analyse our gold standard, urine culture, in the same way as is routinely performed. The alternative suggested by the referee not to use cut-off points used routinely but rather a fixed point is an acceptable but slightly different approach. A third alternative would be the approach used by Juthani-Mehta et al using a combination of positive culture and leukocyturia as the positive gold standard. Unfortunately, switching from present routine culture as the gold standard for any of these alternatives would require collecting new samples from all individuals and redoing the entire study.

#1: Major compulsory revisions: The main problem with the analysis done by the authors is that two different (three) gold standards are used, depending on the presence of symptoms/signs of which the authors state in the background that it is unclear if these symptoms/signs are relevant in bacteriuria (why then letting them influence the gold standard?). As stated above, in my opinion it would be worthwhile (if possible) to analyse the data in two ways.

Our response: We agree with the referee that analysing data several times with slight variations in the definition of the gold standard would be interesting. Unfortunately, we do not have data to support this.

#1: Discretionary revisions: The authors use the word elderly residents. There is some discussion about use of the term 'elderly'. Some suggest the use of 'older adults'. A possible term here could be 'nursing home residents'.

Our response: This is a very interesting point of view. Nursing homes, at least in Sweden, can also include young adults with severe dysfunction. Therefore, we have chosen to specify that we have been studying older adults at nursing homes. In most previously published studies relevant to our study the authors specify the age interval by using the term elderly. Therefore we found it valuable to use the term elderly.

#2. There is very little here of substance. I suggest that the authors beef-up their presentation and spell out the data

Our response: We don't really know what the reviewer means by “beef-up” the presentation. We would be grateful if this could be more specifically explained.

#2. Eliminate the “Yule-Simpson’s” paradox or describe what might be found with this test.

Our response: “Yule-Simpson’s paradox is not a statistical test but a statistical paradox, a phenomenon explaining why the outcome in several groups is changed when groups of different sizes are combined when confounding factors exist. Several
potentially pathogenic bacteria differ in their ability to reduce nitrate to nitrite. Similarly, different bacteria are likely to show a varying ability to provoke pyuria. These differences are confounding factors and as the prevalence of the different types of bacteria varies considerably the size of these groups vary. Thus the outcome of analysing a single bacterium might differ from analysing “any bacteria”. In such cases, results from analysing a single bacterium are more appropriate while results of analysing “any bacteria” are inappropriate. All previously published studies evaluating dipstick urinalysis of the elderly combine different bacterium to “any bacteria” when calculating sensitivity, specificity, PPV or NPV. This has been explained in the background and discussion. For more details of this statistical paradox we refer to the references for this in the manuscript.

#2. Tests were performed on urines (e.g., culture) but no data are given. This reference should be deleted or fixed.

Our response: We have presented data in five tables in the manuscript. Thus, it is not clear to us what the referee means when stating that no data has been presented. It is not clear to us which reference should be deleted or fixed. We would appreciate more information about this.

#2. Accuracy is an elusive and difficult measure. Performing more tests does not improve the accuracy; usually some definitive method is needed to estimate accuracy.

Our response: We don’t understand what the referee means by this comment. We would appreciate if the referee could further explain this.

#2. There are no citations at all, a deficiency. Work like this has been published by others.

Our response: We don’t know what the reviewer means by “there are no citations at all”. We searched literature databases and found about 3000 articles. We then carefully extracted all publications relevant to our study. Finally, we used 30 references in our manuscript. If we have missed important articles on this subject we would be very grateful to know which article/articles we missed.