Research article

Rapid tests and urine sampling techniques for the diagnosis of urinary tract infection (UTI) in children under five years: a systematic review

Penny Whiting[1], Marie Westwood*[2], Ian Watt[3], Julie Cooper[4] and Jos Kleijnen[2]

Address: [1]MRC Health Services Research Collaboration, University of Bristol, England, UK, [2]Centre for Reviews and Dissemination, University of York, England, UK, [3]Department of Health Sciences, University of York, England, UK and [4]Department of Radiology, York District Hospital, England, UK

Email: Penny Whiting - penny.whiting@bristol.ac.uk; Marie Westwood* - mew3@york.ac.uk; Ian Watt - isw1@york.ac.uk; Julie Cooper - Julie.Cooper@york.nhs.uk; Jos Kleijnen - jk13@york.ac.uk

* Corresponding author

Abstract

Background: Urinary tract infection (UTI) is one of the most common sources of infection in children under five. Prompt diagnosis and treatment is important to reduce the risk of renal scarring. Rapid, cost-effective, methods of UTI diagnosis are required as an alternative to culture.

Methods: We conducted a systematic review to determine the diagnostic accuracy of rapid tests for detecting UTI in children under five years of age.

Results: The evidence supports the use of dipstick positive for both leukocyte esterase and nitrite (pooled LR+ = 28.2, 95% CI: 17.3, 46.0) or microscopy positive for both pyuria and bacteriuria (pooled LR+ = 37.0, 95% CI: 11.0, 125.9) to rule in UTI. Similarly dipstick negative for both LE and nitrite (Pooled LR- = 0.20, 95% CI: 0.16, 0.26) or microscopy negative for both pyuria and bacteriuria (Pooled LR- = 0.11, 95% CI: 0.05, 0.23) can be used to rule out UTI. A test for glucose showed promise in potty-trained children. However, all studies were over 30 years old. Further evaluation of this test may be useful.

Conclusion: Dipstick negative for both LE and nitrite or microscopic analysis negative for both pyuria and bacteriuria of a clean voided urine, bag, or nappy/pad specimen may reasonably be used to rule out UTI. These patients can then reasonably be excluded from further investigation, without the need for confirmatory culture. Similarly, combinations of positive tests could be used to rule in UTI, and trigger further investigation.

Background

Urinary tract infection (UTI) is one of the most common sources of infection in children under 5. In a small proportion of children UTI may lead to renal scarring [1,2].

This outcome of infection is of concern as it is associated with significant future complications and ultimately with end stage renal disease[3]. Prompt diagnosis and treatment is therefore important to reduce the risk of future scarring.

Table 1: Details of tests evaluated in the review

Test	Details	Advantages	Disadvantages
Urine sampling			
Suprapubic aspiration (SPA)	Needle attached to syringe inserted through lower abdomen into bladder.	Least risk of contamination	Invasive
Transurethral catheterisation	Catheter inserted through the urethra into the bladder.	Less invasive than SPA	Invasive, causes pain and distress to child
Clean voided urine (CVU)	Midstream sample collected in sterile container.	Non-invasive, easy to obtain	Difficult in younger children
Urine bags	Bag applied to perineum.	Suitable for babies and infants	Risk of contamination
Urine pads	Absorbent pad placed in nappy.		
Dipstick			
Nitrite	Gram-negative bacteria reduce dietary nitrate to nitrites.	Very easy and quick to perform, relatively cheap	Less accurate than culture
Leukocyte esterase (LE) Glucose	Leukocyte esterase is an enzyme that suggests the presence of leukocytes. Normal urine contains small amount of glucose. Bacteria metabolise glucose and so this test tests for the absence of glucose. Requires morning fasting urine specimen.		Not commercially available, not suitable for non-potty trained children
Microscopy			
Pyuria	Urine examined through microscope for presence of white blood cells. Samples may be centrifuged before examination	Quicker than culture	More time consuming than dipstick, more expensive than dipstick and culture
Bacteriuria	Urine examined for presence of bacteria. Urine may be Gram-stained.		
Culture			
Standard Culture	Reference standard test for UTI. Involves streaking urine on enrichment and selective media.	Very accurate	Time consuming: takes 48 hours to give a result, has to be performed in the laboratory

Clinical history and examination is the first step in any diagnosis and is the means of identifying children with suspected UTI. Elements of the clinical examination have also been evaluated as diagnostic tests for UTI but there is little data available on these. Urine tests are commonly used for the diagnosis of UTI.

The reference standard for the diagnosis of UTI in children is considered to be any bacterial growth on a culture of urine obtained by suprapubic aspiration[4]. Culture has the disadvantage of taking at least 48 hours to give a result. More rapid methods of UTI diagnosis are therefore desirable. The most widely used rapid tests are dipsticks.

Analytes commonly tested by dipsticks include leukocyte esterase, nitrite, blood and protein[4]. Dipstick tests have the advantage of being quick and easy to perform and can be carried out in primary care giving an immediate result. Microscopic examination of urine samples for leukocytes or bacteria [4] is considerably more time consuming and labour intensive than the dipstick method[5]. However, unlike culture, it can be used to give results within the primary care setting. An uncontaminated sample is necessary to reach an accurate diagnosis. Obtaining this is a particular issue when investigating young children. Table 1 presents a summary of the advantages and disadvantages of these tests.

A wide range of other tests have been evaluated for the diagnosis of UTI. These include dipslide and rapid culture methods, colorimetric tests, headspace gas analysis, impedance, bio- and chemical luminescence, immunologic tests (e.g. ELISA), enzyme tests, bacterial oxygen consumption, and turbidimetry. However, these are not in widespread use and will not be discussed in this paper.

This review aims to determine the diagnostic accuracy of dipstick and microscopy, and different methods of urine sampling, for detecting UTI in children under five years of age. Two previous reviews have addressed a similar objective[6,7]. These were published over 2 years ago and did not assess urine sampling. They also included fewer studies (48 and 26 compared to 70), possibly as a result of less extensive literature searches and tighter inclusion criteria, than this review. This review therefore presents the most up to date and extensive systematic review of the topic area.

Methods

We searched 16 electronic databases from inception to between October 2002 and February 2003. Update searches were conducted in May 2004. To identify additional published and unpublished studies we searched the internet, hand searched 12 key journals, screened reference lists of included papers and contacted experts in the field.

We did not apply any language restrictions. Full details of the search strategy will be reported elsewhere[8].

Studies had to meet the following criteria to be included in the review:

Study design: diagnostic cohort (single sample) studies

Population: at least some children aged <5 years with suspected UTI

Index tests: microscopy or dipstick tests used to diagnose UTI or an evaluation of urine sampling methods.

Reference standard: culture or culture combined with other tests

Outcome measures: sufficient information to construct a 2 × 2 table

Two reviewers independently screened titles and abstracts for relevance, we resolved disagreements by consensus. One reviewer performed inclusion assessment; data extraction and quality assessment and a second reviewer checked this. We extracted 2 × 2 data and used this to calculate measures of diagnostic performance. We used QUADAS to assess study quality[9]. Individual QUADAS items were used to investigate heterogeneity and to present a detailed assessment of quality to the reader.

For each test, or test combination, we calculated the range in sensitivity, specificity, positive (LR+) and negative (LR-) likelihood ratios, and diagnostic odds ratios (DOR). We selected likelihood ratios as the measure of test performance for further analysis as these measures are easier to interpret than sensitivity and specificity [10]. For tests investigated in more than two studies, we used random effects models to pool positive and negative likelihood ratios [11]. Where studies presented more than one estimate of test performance for the same test, for example at different cut-off points or for different patient subgroups, we only included one estimate in the pooled analysis. We aimed to select the data set most similar to the estimates provided by the other studies in terms of population, test manufacturer or population. Heterogeneity of likelihood ratios was investigated using the Q statistic [12] and through visual examination of forest plots of study results [13].

We presented individual studies results graphically by plotting estimates of sensitivity and specificity in receiver operating characteristic (ROC) space. Where sufficient data were available, we used regression analysis to investigate heterogeneity. We extended the summary ROC (sROC) model [14], estimated by regressing D (log DOR) against S (logit true positive rate – logit false positive rate), weighted according to sample size, to include covariates relating to patient age (<2 years, <5 years, <12 years and <18 years), geographic region and each of the 14 QUADAS items. In addition, for microscopy for pyuria and bacteriuria a variable on whether the sample was centrifuged was included, and for microscopy for bacteriuria a variable for Gram stain was included.

Results

The literature searches identified over 10 000 references of which 70 studies were included. Figure 1 shows the flow of studies through the review process. A summary of the results of all 70 studies included in the review is provided [see Additional file 1].

Quality

The median number of the 14 QUADAS items fulfilled was 8 (range 5–13). The main limitation with the studies was the failure to include an appropriate patient spectrum (<40%) or to report inclusion criteria. Studies also failed to report sufficient details to judge whether clinical review bias (the availability of clinical information to the person interpreting the test results), diagnostic review bias (the availability of the results of the index test to the person interpreting the reference standard) and test review bias
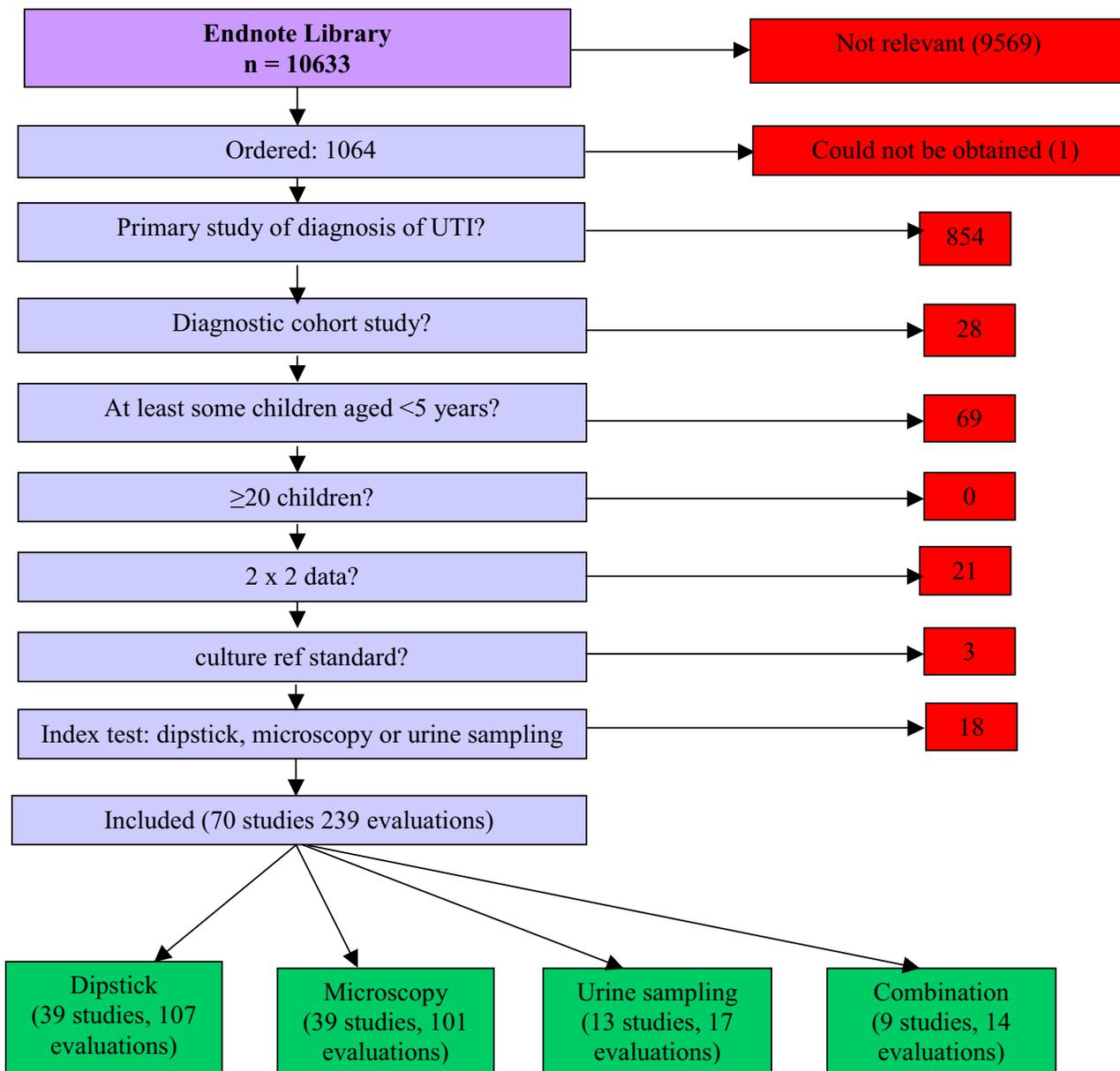
Figure 1
Flow chart of studies through review process.

(the availability of the results of the reference standard to the person interpreting the index test) were avoided. Withdrawals and handling of uninterpretable results were also poorly reported. Figure 2 illustrates the number of studies that answered "yes", "no" and "not stated" to each of the 14 QUADAS items. A summary of the results of the quality assessment for each study is provided [see 1].

Urine sampling
Thirteen studies, with a total of 17 different test evaluations, compared the diagnostic accuracy of different methods obtaining urine for testing [15-27]. These studies compared the results of culture from urine obtained by different sampling methods. Five studies reporting seven data sets assessed the diagnostic accuracy of a clean voided
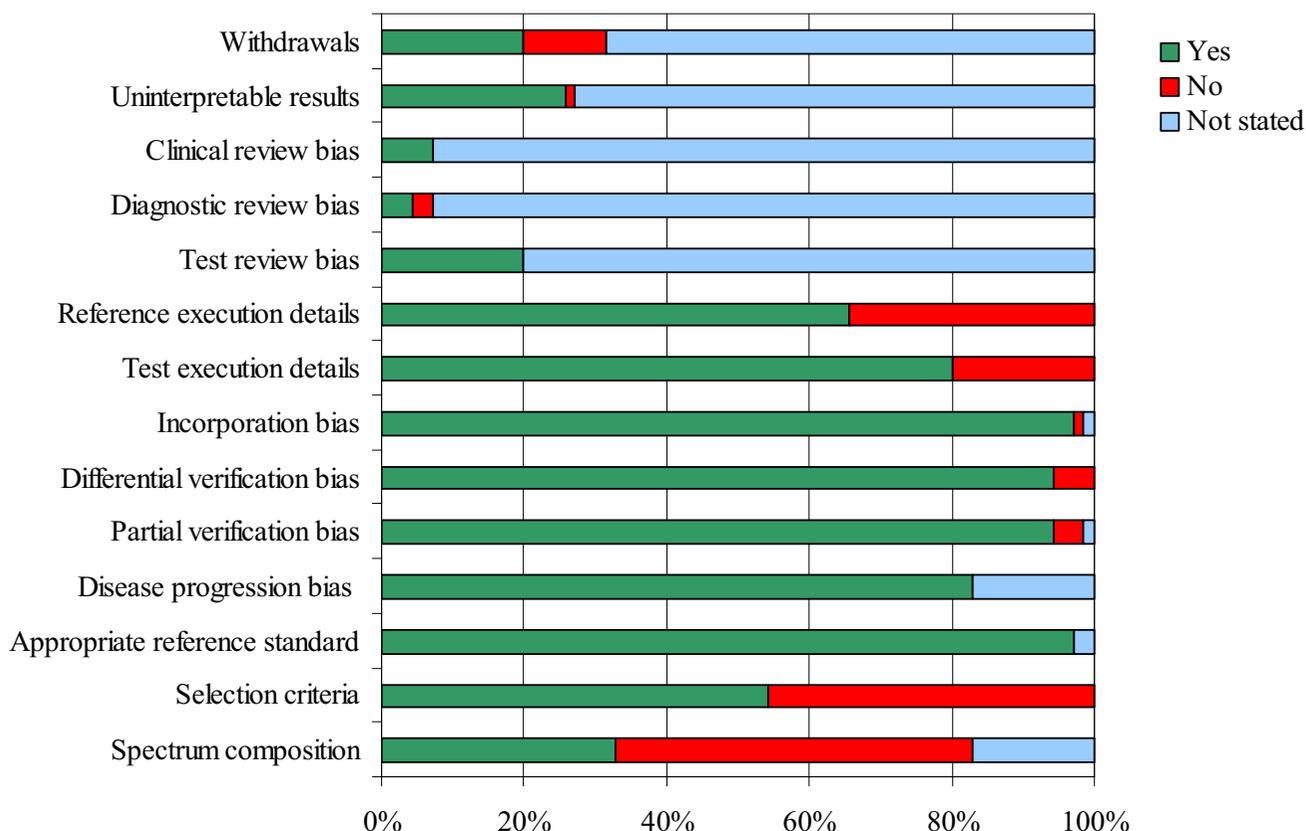
Figure 2
Results of the quality assessment.

Table 2: Summary of results for studies of dipstick tests

Dipstick positive for:	Number of studies	Range in LR+	Pooled LR+ (95 % CI)*	Range in LR-	Pooled LR- (95 % CI)*
Nitrite	23	2.5 – 439.6	15.9 (10.7, 23.7)	0.12 – 0.86	0.51 (0.43, 0.60)
LE	14	2.6 – 32.2	5.5 (4.1, 7.3)	0.02 – 0.66	0.26 (0.18, 0.36)
Nitrite or LE	15	3.0 – 32.2	6.1 (4.3, 8.6)	0.03 – 0.39	0.20 (0.16, 0.26)
Nitrite and LE	9	6.3 – 197.1	28.2 (17.3–46.0)	0.07 – 0.86	0.37 (0.26, 0.52)
Glucose	4	25.2 – 156.1	66.3 (20.0, 219.6)	0.02 – 0.38	0.07 (0.01, 0.83)
Protein	2	1.7 & 1.8	na	0.78 & 0.96	na
Blood	1	2.3	na	0.84	na
LE and protein	1	17.4	na	0.12	na
Nitrite, blood, or protein	1	2.7	na	0.28	na
Nitrite, blood, or LE	1	1.3	na	0.50	na
Nitite, blood and LE	1	3.5	na	0.19	na
Nitrite, LE and protein	2	3.1 & 69.2	na	0.05 & 0.17	na
Nitrite, LE, or protein	1	1.9	na	0.05	na
Nitrite, LE, protein, or blood	1	8.0	na	0.19	na

* There was significant heterogeneity in all pooled estimates therefore these should be interpreted with caution

urine (CVU) sample, using a supra-pubic aspiration (SPA) urine sample as the reference standard [15-19]. When both samples were cultured the agreement between the two sampling methods was good. There was considerable heterogeneity in positive likelihood ratios (p < 0.0001). However, the negative likelihood ratios were statistically homogeneous (p = 0.531). The pooled positive likelihood ratio for a CVU sample was 8.8 (95% CI: 2.6, 29.6) and the pooled negative likelihood ratio was 0.23 (95% CI: 0.18, 0.30). Overall, there were insufficient data to draw any conclusions regarding the appropriateness of using urine samples obtained from bags (4 studies)[16,20,21,27] or pads/nappies (4 studies) [22-25].

Dipstick tests

A total of 39 studies reporting 107 data sets evaluated dipstick tests for the diagnosis of UTI [28-66]. These studies assessed the utility of dipstick tests for nitrite, leukocyte esterase (LE), protein, glucose and blood, alone and in combination. Table 2 summarises the results of these studies.

Figure 3 shows the estimates of sensitivity and 1-specificity plotted in ROC space for glucose, and dipstick tests for nitrite and LE, alone and in combination. This graph suggests that glucose is considerably better than the other tests, both for ruling in and ruling out disease, this is supported by the pooled likelihood ratios. However, the confidence intervals around the pooled likelihood ratios are very large, especially for the negative likelihood ratios (ruling out disease), suggesting considerable uncertainty in these estimates. It should also be noted that very few studies of glucose tests were available and that they were all conducted over 30 years ago and the test used ("Uriglox") [58] is no longer commercially available.

Nitrite alone has a relatively high pooled positive likelihood ratio (15.9, 95% CI: 10.7, 23.7) and so may be useful for ruling in disease. However, it has a relatively poor negative likelihood ratio (0.51, 95% CI: 0.43, 0.60) suggesting that it may not be a useful test for ruling out disease. LE alone appears to be a relatively poor test both for ruling in (pooled LR+ = 5.5, 95% CI: 4.1, 7.3) and ruling out disease (pooled LR- = 0.26, 95% CI: 0.18, 0.36). A strategy which combines the results of LE and nitrite testing appears to offer the best performance both for ruling in and ruling out disease. A dipstick test positive for both nitrite and LE has the highest positive likelihood ratio (28.2, 95% CI: 17.3, 46.0) suggesting that this test combination may be used to rule in disease. A dipstick test negative for both LE and nitrite has the best negative likelihood ratio (0.20, 95% CI: 0.16, 0.26) suggesting that this test combination may be used to rule out disease. A dipstick test positive for either LE or nitrite and negative
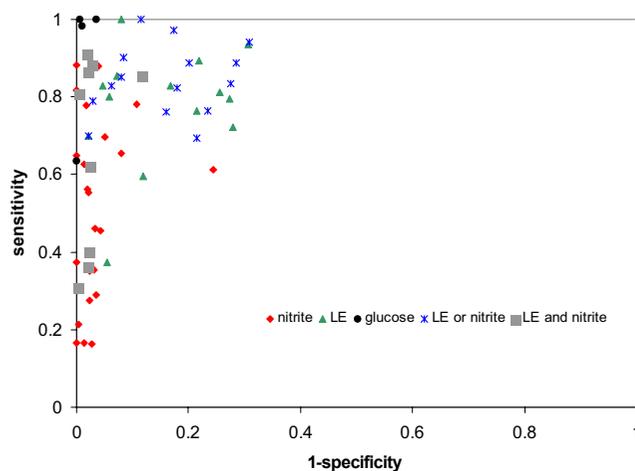

Figure 3
Sensitivity and specificity plotted in ROC space for different dipstick tests.

for the other is less informative for the diagnosis of UTI. Such a test result could be seen as an "indeterminate" test result requiring further investigation.

It is difficult to draw conclusions about the overall accuracy of dipstick tests given the heterogeneity between studies in some areas, and the lack of data in others. There was insufficient information to make any judgement regarding the overall diagnostic accuracy of dipstick tests for protein, blood, or for combinations of three different dipstick tests (e.g. combination of LE, nitrite and blood).

A regression analysis found that only clinical review bias showed an association with the diagnostic accuracy of nitrite dipstick (the DOR was 3.1 (95% CI: 0.97, 9.95) times higher in studies that avoided clinical review bias, i.e. in those studies that reported that the same clinical information was available to those interpreting the test results as would be available in practice). A higher DOR indicates higher overall accuracy. None of the items investigated, including age, showed a significant association with the DOR in the regression analysis for dipstick for LE, or for dipstick for LE or nitrite positive. Regression analysis was not carried out to investigate heterogeneity for other tests, as insufficient data were available.

Microscopy

A total of 39 studies reporting 101 data sets evaluated microscopy for diagnosing UTI [15-18,26,28-33,35,36,38,42,52,59-61,64,67-84]. Microscopy was used

Table 3: Summary of results for studies of microscopy

Microscopy positive for:	Number of studies	Range in LR+	Pooled LR+ (95 % CI)*	Range in LR-	Pooled LR- (95 % CI)*
Pyuria	28	1.3 – 27.7	5.9 (4.1, 8.5)	0.04 – 0.68	0.27 (0.20, 0.37)
Bacteriuria	22	1.6 – 304.8	14.7 (8.6, 24.9)	0.01 – 0.48	0.19 (0.14, 0.24)
Pyuria or bacteriuria	8	1.5 – 5.9	4.2 (2.3, 7.6)	0.02 – 0.27	0.11 (0.05, 0.23)
Pyuria and bacteriuria	8	2.7 – 281.0	37.0 (11.0, 125.9)	0.07 – 0.56	0.21 (0.13, 0.36)

* There was significant heterogeneity in all pooled estimates therefore these should be interpreted with caution
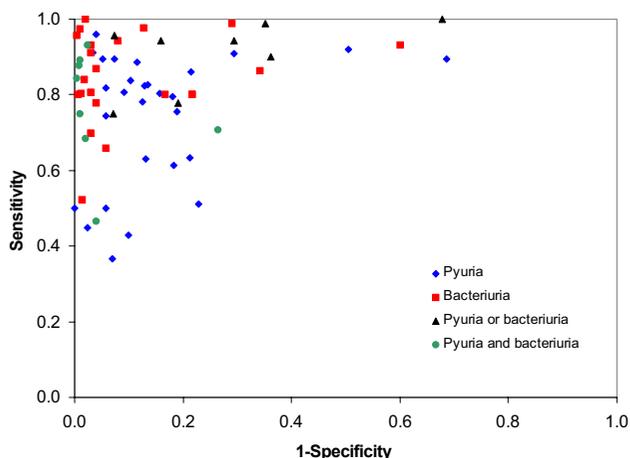

Figure 4
Sensitivity and specificity plotted in ROC space for different microscopy evaluations.

to determine the presence of pyuria or bacteriuria, or combinations of the two. Table 3 summarises the results of these studies.

Figure 4 shows the estimates of sensitivity and 1-specificity plotted in ROC space for all studies. This graph suggests that bacteriuria is considerably better than pyuria both for ruling out and ruling in disease. The diagnostic performance of bacteriuria may be improved when combined with pyuria. The pooled positive likelihood ratios are highest for pyuria and bacteriuria combined (37.0, 95% CI: 11.0, 125.9), where a positive result was defined as both tests positive, supporting the suggestion that the combination of a positive result for both of these tests may be useful for ruling in disease. Conversely, the lowest negative likelihood ratio resulted from the combination of pyuria and bacteriuria (0.11, 95% CI: 0.14, 0.24), where a negative result was defined as both tests negative, and this may be useful for ruling out disease.

However, the confidence intervals around the pooled estimates are large, and in combination with the observed heterogeneity suggest considerable uncertainty in these estimates.

Regression analysis showed that centrifugation of the sample, reporting of selection criteria, reporting of details of reference standard execution, and reporting of uninterpretable results showed a significant association with the DOR in the studies of microscopy for pyuria. All of these items, with the exception of centrifugation, relate to the quality of reporting. The DOR was 6.25 (95% CI: 3.44, 11.11) times greater in samples that were not centrifuged; 3.19 (95% CI: 1.76, 5.79) times higher in studies that adequately reported selection criteria; 6.6 (95% CI: 2.43, 17.94) times higher in studies that reported sufficient details of reference standard execution; and 2.99 (95% CI: 1.50, 5.94) times higher in studies that reported on uninterpretable results. The association for centrifugation is not what we anticipated, as we would expect centrifugation of the sample to lead to improved test accuracy.

In the analysis of microscopy for bacteriuria, Gram stain, incorporation bias and reporting of selection criteria showed a significant association with the DOR. The DOR was 5.96 (95% CI: 2.99, 11.89) times greater in samples that were Gram stained; 50.0 (95% CI: 6.67, 1000) times greater in studies in which incorporation bias was not present (i.e. studies in which the index test did not form part of the reference standard); and 2.46 (95% CI: 1.26, 8.41) times greater in studies that reported selection criteria. We would expect Gram staining to increase test performance as found in the analysis. However, we would expect the absence of incorporation bias to decrease test performance. The observed association may be explained by the fact that, for the purposes of the regression analysis, studies scoring "unclear" for a quality item were grouped with those scoring "no". For this analysis only one study scored "no" (i.e. incorporation bias was present) and the other studies grouped with this scored as "unclear". The association may therefore reflect quality of reporting, as does the association with reporting of selection criteria.

Combinations of tests from different categories

Nine studies including a total of 14 data sets examined the accuracy of different combinations of microscopy and dipstick tests for the diagnosis of UTI [30,32,35,36,42,52,67,79,85]. Given the results of individual tests, the test combination that appears to be potentially the most interesting is dipstick for LE and nitrite, and microscopy for pyuria and bacteriuria. Five studies investigated different permutations of these tests. [32,35,42,67,79] Three studies evaluated the accuracy of a positive result in one of these four tests (i.e. dipstick positive for LE or nitrite or microscopy positive for pyuria or bacteriuria) [32,35,67]. The results varied considerably between studies with positive likelihood ratios ranging from 0.8 to 35.9, and negative likelihood ratios ranging from 0.01 to 5.38. It is therefore not possible to draw overall conclusions from these studies. One study examined the combination of a positive result for all four tests [35]. This study reported a very high positive likelihood ratio (35.9) i.e. the combination was found to be very good for ruling in disease, but the negative likelihood ratio was less good at 0.28. These results might be expected given the results from the studies that examined combinations of dipstick tests, or combinations of microscopy tests.

The other test combinations evaluated by these studies differed widely, and none were repeated between studies. Test combinations investigated included LE and nitrite dipstick test combined with microscopy for bacteriuria[42] or pyuria [42,52,85], dipstick for LE, nitrite and blood combined with microscopy for pyuria,[36] and dipstick for nitrite combined with microscopy for pyuria[30].

As most test combinations were only evaluated by one study and the definition of a positive test varied for the tests investigated by more than one study, it was not possible to draw conclusions regarding the diagnostic accuracy of these test combinations.

Comparison of different tests

Comparison of the pooled likelihood ratios suggests that the microscopy combinations may be more accurate than the dipstick combinations. Only one study evaluated both dipstick positive for nitrite and LE and microscopy positive for bacteriuria and pyuria [59]. This study found that the dipstick combination was best for ruling in disease (LR+ was 18.9 for the dipstick combination compared to 11.6 for the microscopy combination). Five studies examined dipstick negative for nitrite and LE, and microscopy negative for pyuria and bacteriuria [32,35,40,42,59]. All but one found that microscopy was better for ruling out disease than dipstick.

What do these results mean?

If we take an estimate for the prevalence of UTI in children presenting to their GP with symptoms of possible UTI (the pre-test probability of disease), i.e. children in whom tests to diagnose UTI are likely to be used, likelihood ratios can be used to calculate the post-test probability of UTI. We were unable to find reliable estimates of the pre-test probability of UTI in the literature, and therefore used the results from the included studies to provide an estimate. Only studies that included an appropriate patient spectrum were included in this analysis. UTI prevalence varied greatly between studies (3–73%). As the distribution was highly skewed we used the median prevalence, which was 20%. Figure 5 shows how the probability of UTI changes after testing. In a typical primary care setting in which the pre-test probability of disease is estimated to be around 20%, a negative likelihood ratio of 0.20 translates to a post-test probability of UTI of about 4%. In other words, children who receive a dipstick test negative for both nitrite and LE have a 4% probability of having a UTI.

Discussion

An accurate and prompt diagnosis is important to inform patient management decisions in young children with suspected UTI. The first step in the diagnostic process is to identify children presenting to the GPs surgery who may have a UTI. This will inevitably involve a clinical assessment. It is very difficult, if not impossible, to capture all the signs and symptoms that a GP might use to develop a clinical suspicion of UTI and decide to test a child for UTI. Further research to accurately define from which children urine samples should be taken to test for UTI may be useful.

Following clinical examination, the next step is to collect a suitable urine sample to test for the presence of infection. Different methods of urine sampling may be differently susceptible to contamination and hence to false positive results. The issue of appropriate urine sampling techniques is of particular concern in young children, where the collection of a sterile, mid-stream sample can be problematic. Suprapubic aspiration has been regarded as the reference standard collection method. This procedure is invasive and may require the use of ultrasound guidance to ensure that the needle is inserted into the bladder. The identification of an alternative sampling method with acceptable diagnostic performance, which can readily be applied in the GP's surgery, and which is more acceptable to children and parents, is therefore desirable. The studies on urine sampling showed reasonably good agreement between clean voided urine (CVU) and suprapubic aspiration (SPA) samples, suggesting that this is an appropriate routine method of urine collection. CVU samples are difficult to collect in young children
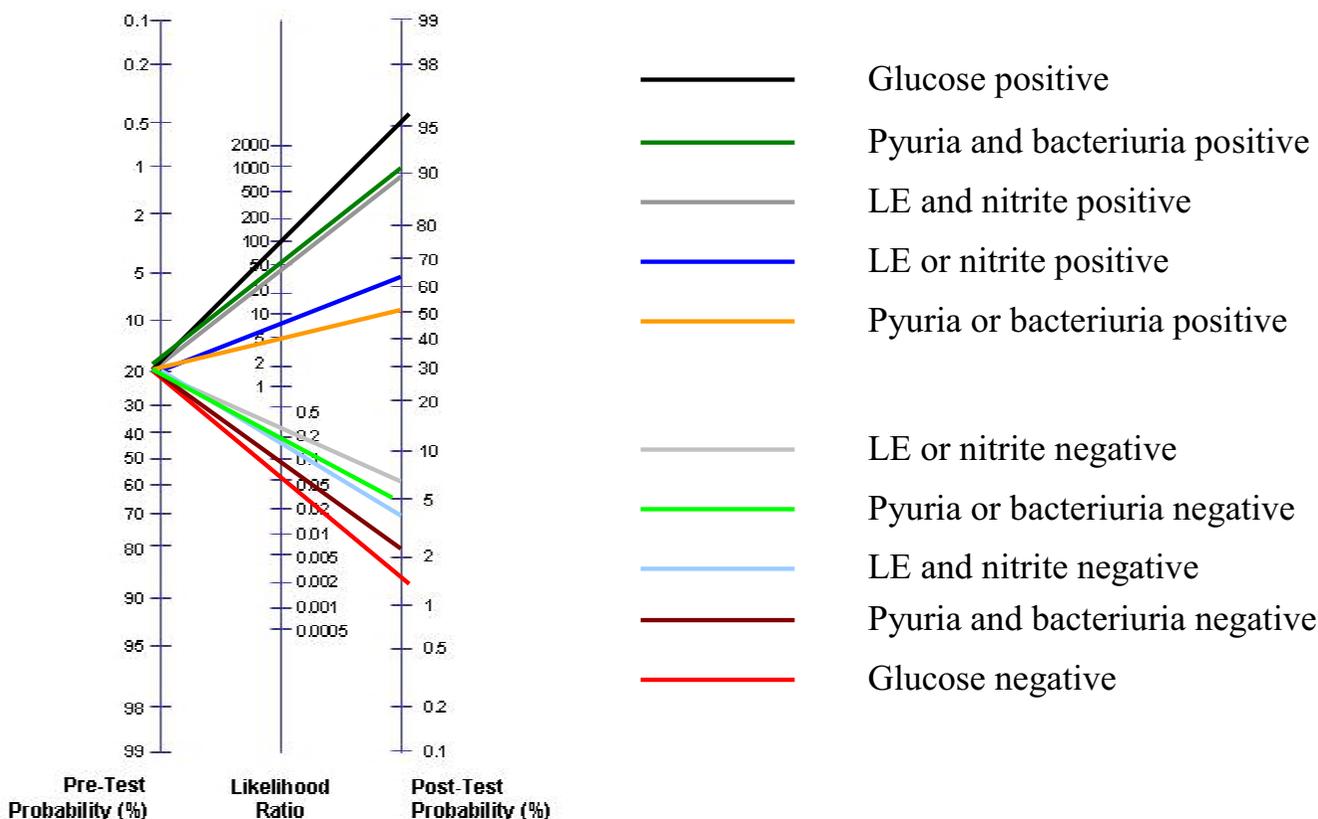
Figure 5
Likelihood ratio nomogram for dipstick tests.

who are not potty trained. A number of alternative collection methods have been developed, including bag, pad and nappy specimens. There is currently insufficient data available to determine whether bag or nappy/pad specimens may be used as substitutes for SPA Further work is needed in this area.

The main types of urine testing evaluated for the diagnosis of UTI were dipstick and microscopy. Culture is generally considered to be the reference standard for UTI diagnosis. The logistics of urine culture represent a significant drawback; culture takes approximately 48 hours to give a result, is generally performed in the laboratory and is more expensive than other methods. For this reason alternative, more rapid tests are needed to guide the prompt initiation of treatment. Dipsticks have the advantage of providing an immediate result, and of being both cheap and easy to perform and interpret. The studies of dipstick tests showed considerable heterogeneity and so the results

should be interpreted with caution. The results suggest that a dipstick test that is positive for both LE and nitrite is good for ruling in disease whilst one that is negative for both LE and nitrite is good for ruling out disease.

An additional dipstick test that provided interesting results was the estimation of urinary glucose, where a negative urinary glucose is regarded as a positive test for UTI. Only four studies of this test were identified, and all were conducted more than 30 years ago. All studies reported excellent specificity for this test. Sensitivity was also very high in three of the studies but was lower, at 64% in the fourth. This last study was conducted in children aged less than one year, suggesting that the test maybe less useful in very young children. This difference in performance of the test with patient age may be explained by its apparent dependence on an overnight, fasting sample; such a sample would be impossible to obtain in children who are not toilet trained. However, given the limited results reported,

this test appears to be potentially useful for the diagnosis of UTI in toilet trained children. Further studies are needed.

Although, in practice, microscopy and culture are generally requested in combination, microscopy has the advantage of being quicker to provide a result. It may be that microscopy has some potential as a test that could be performed in the GP surgery. However, it remains more expensive than a dipstick test and requires some degree of expertise to perform. The studies of microscopy showed considerable heterogeneity, in terms of results, cut-off points, types of urine samples and population. A urine sample that was positive for both pyuria and bacteriuria on microscopy was found to be very good for ruling in disease. Similarly, a urine sample that was negative for both pyuria and bacteriuria on microscopy was found to be very good for ruling out disease.

The possibility of publication bias remains a potential problem in this review. It is possible, and indeed likely, that studies reporting higher estimates of test performance are more often published, but the extent to which this occurs is unclear. There is evidence that publication bias is a particular problem for studies of small sample size, although these data are general and does not come from the diagnostic literature[86,87]. We restricted this review to studies that included at least 20 children, meaning that this type of publication bias is less likely to be a problem. We are unaware of any articles on publication bias in diagnostic tests or on methods to formally assess publication bias in a diagnostic systematic review.

We chose likelihood ratios as the primary effect measure as these are the measure that physicians find easiest to interpret [88]. We used pooled likelihood ratios and estimates of the pre-test probability of disease to calculate estimates of the post-test probability of disease. These measures provide a simple illustration of how the results of a test change the probability of disease and help the reader to determine how useful a test is likely to be in practice. The main limitation of this approach was the considerable heterogeneity in pooled likelihood ratios; it is debatable whether it is appropriate to pool these estimates. It is important that pooled estimates are interpreted with caution and that the heterogeneity between studies is considered when interpreting these results. A further problem with this analysis is that positive and negative likelihood ratios were pooled individually. These measures are likely to be correlated within an individual study and ignoring this correlation may be problematic[89].

We conducted a regression analysis to investigate possible explanations for the observed heterogeneity. This analysis was carried out according to standard methods for pooling studies of diagnostic accuracy using the summary ROC approach[14]. Using the DOR for further investigation of heterogeneity means that we can only assess whether the factors investigated are associated with the DOR and not with sensitivity and specificity, or with positive and negative likelihood ratios. Often factors that lead to an increase in sensitivity will lead to a decrease in specificity and vice versa, possibly with no effect on the DOR. A further limitation of this analysis was that we could only investigate the effect of variables at the study level. One factor that may impact on the accuracy of the diagnostic tests investigated is patient age. However, as the majority of studies investigated included children aged 0–16 or 18 years and did not report results separately for younger age groups it was not possible to carry out appropriate subgroup analyses to investigate the effects of age on estimates of test accuracy. This is an area where further investigation is required.

Conclusion

Based on the results of this review dipstick negative for LE and nitrite, or microscopic analysis negative for pyuria and bacteriuria of a CVU, bag, or nappy/pad specimen may reasonably be used to rule out UTI. These patients can then be excluded from further investigation, without the need for confirmatory culture. Similarly, combinations of positive tests could be used to rule in UTI, and trigger further investigation. In the latter case, however, confirmation by culture may be preferred prior to the initiation of further, possibly invasive, investigations. Additional information on antibiotic sensitivities, which can be provided by culture, may also be a significant consideration. If combinations of rapid tests were routinely used to rule in and/or rule out disease, as described, then a cost saving in the number of cultures ordered would be expected. In addition it is likely that the number of children without disease exposed to inappropriate antibiotic therapy, whilst awaiting culture results, would be reduced. This may have implications for antibiotic resistance at a population level.

The quality assessment highlighted several areas that could be improved upon in future diagnostic accuracy studies, in particular in relation to reporting. Future studies should follow the STARD guidelines for reporting of diagnostic accuracy studies [90].

The review also highlighted the following specific areas requiring further research for the diagnosis of UTI:

• urine sampling methods in younger children

• accuracy of the glucose test, and its practical applicability

• handling of indeterminate nitrite and LE dipstick test results

• accuracy of microscopy in combination with a dipstick test

Abbreviations
CI Confidence interval

CVU Clean voided urine

DOR Diagnostic odds ratio

ELISA Enzyme linked immunosorbant assay

LE Leukocyte esterase

LR Likelihood ratio

QUADAS Quality assessment of diagnostic accuracy studies

ROC Receiver operating characteristic

SPA Suprapubic aspiration

STARD Standards of reporting of diagnostic accuracy studies

UTI Urinary tract infection

Competing interests
The author(s) declare that they have no competing interests.

Authors' contributions
All authors contributed towards the conception and design of the study and the interpretation of the data. They also read and approved the final manuscript. PW and MW participated in data extraction, the analysis of data, and drafted the article.

Additional material

Acknowledgements

References
1. **The management of urinary tract infection in children.** *DRUG AND THERAPEUTICS BULLETIN* 1997, **35**:65-69.
2. Vernon SJ, Coulthard MG, Lambert HJ, Keir MJ, Matthews JN: **New renal scarring in children who at age 3 and 4 years had had normal scans with dimercaptosuccinic acid: follow up study.** *Bmj* 1997, **315**:905-908.
3. Larcombe J: **Urinary tract infection.** In *Clinical Evidence, Issue 7 Volume 7.* London, BMJ Publishing; 2002:377-385.
4. Downs SM: **Technical report: urinary tract infections in febrile infants and young children. The Urinary Tract Subcommittee of the American Academy of Pediatrics Committee on Quality Improvement.** *Pediatrics* 1999, **103**:e54.
5. Upadhyay J, Bolduc S, Braga L, Farhat W, Bagli DJ, McLorie GA, Khoury AE, El-Ghoneimi A: **Impact of prenatal diagnosis on the morbidity associated with ureterocele management.** *Journal of Urology* 2002, **167**:2560-2565.
6. Huicho L, Campos-Sanchez M, Alamo C: **Metaanalysis of urine screening tests for determining the risk of urinary tract infection in children.** *Pediatric Infectious Disease Journal* 2002, **21**:1-11, 88.
7. Gorelick MH, Shaw KN: **Screening tests for urinary tract infection in children: A meta-analysis.** *Pediatrics* 1999, **104**:e54.
8. Whiting P, Westwood M, Ginnelly L, Palmer S, Richardson G, Cooper J, Watt I, Glanville J, Sculpher M, Kleijnen J: **A systematic review of tests for the diagnosis and evaluation of urinary tract infection (UTI) in children under five years.** *HEALTH TECHNOLOGY ASSESSMENT* in press.
9. Whiting P, Rutjes AWS, Reitsma JB, Bossuyt PM, Kleijnen J: **The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews.** *BMC Medical Research Methodology* 2003, **3**:1-25.
10. Steurer J, Fischer JE, Bachmann LM, Koller M, ter Riet G: **Communicating test accuracy terms to practicing physicians - a controlled study.** *BMJ* 2002, **324**:824.

11. DerSimonian R, Laird N: **Meta-analysis in clinical trials.** *Controlled Clinical Trials* 1986, **7:**177-188.
12. Fleiss JL: **The statistical basis of meta-analysis.** *Statistical Methods in Medical Research* 1993, **2:**121-145.
13. Galbraith RF: **A note on graphical representation of estimated odds ratios from several clinical trials.** *Stat Med* 1988, **7:**889-894.
14. Moses LE, Shapiro D, Littenberg B: **Combining independent studies of a diagnostic test into a summary ROC curve: data-analystic approaches and some additional considerations.** *Statistics in Medicine* 1993, **12:**1293-1316.
15. Aronson AS, Gustafson B, Svenningsen NW: **Combined suprapubic aspiration and clean-voided urine examination in infants and children.** *Acta Paediatrica Scandinavica* 1973, **62:**396-400.
16. Hardy JD, Furnell PM, Brumfitt W: **Comparison of sterile bag, clean catch and suprapubic aspiration in the diagnosis of urinary infection in early childhood.** *British Journal of Urology* 1976, **48:**279-283.
17. Morton RE, Lawande R: **The diagnosis of urinary tract infection: comparison of urine culture from suprapubic aspiration and midstream collection in a children's out-patient department in Nigeria.** *Annals of tropical paediatrics* 1982, **2:**109-112.
18. Pylkkanen J, Vilska J, Koskimies O: **Diagnostic value of symptoms and clean-voided urine specimen in childhood urinary tract infection.** *Acta Paediatrica Scandinavica* 1979, **68:**341-344.
19. Ramage IJ, Chapman JP, Hollman AS, Elabassi M, McColl JH, Beattie TJ: **Accuracy of clean-catch urine collection in infancy.** *Journal of Pediatrics* 1999, **135:**765-767.
20. Braude H, Forfar JO, Gould JC, McLeod JW: **Diagnosis of urinary tract infection in childhood based on examination of pared non-catheter and catheter specimens of urine.** *British Medical Journal* 1967, **4:**702-705.
21. Benito Fernandez J, Sanchez Echaniz J, Mintegui Raso S, Montejo F: **Urinary tract infection in infants: Use of urine specimens obtained by suprapubic bladder aspiration in order to determine the reliability of culture specimen of urine collected in perineal bag.** *Anales Espanoles de Pediatria* 1996, **45:**149-152.
22. Farrell M, Devine K, Lancaster G, Judd B: **A method comparison study to assess the reliability of urine collection pads as a means of obtaining urine specimens from non-toilet-trained children for microbiological examination.** *Journal of Advanced Nursing* 2002, **37:**387-393.
23. Feasey S: **Are Newcastle urine collection pads suitable as a means of collecting specimens from infants?** *Paediatric Nursing* 1999, **11:**17-21.
24. Ahmad T, Vickers D, Campbell S, Coulthard MG, Pedler S: **Urine collection from disposable nappies.** *Lancet* 1991, **338:**674-676.
25. Cohen HA, Woloch B, Linder N, Vardi A, Barzilai A: **Urine samples from disposable diapers: an accurate method for urine cultures.** *Journal of Family Practice* 1997, **44:**290-292.
26. Dayan PS, Chamberlain JM, Boenning D, Adirim T, Schor JA, Klein BL: **A comparison of the initial to the later stream urine in children catheterized to evaluate for a urinary tract infection.** *Pediatric Emergency Care* 2000, **16:**88-90.
27. Mendez EB: **Are positive urine cultures obtained using recollector bags reliable?** *Revista Chilena de Pediatria* 2003, **74:**487-491.
28. Dayan PS, Bennett J, Best R, Bregstein JS, Levine D, Novick MK, Sonnett FM, Stimell-Rauch ML, Urtecho J, Wagh A, Miller SZ: **Test characteristics of the urine Gram stain in infants less than 60 or 60 days of age with fever.** *Pediatric Emergency Care* 2002, **18:**12-14.
29. Armengol CE, Hendley JO, Schlager TA: **Should we abandon standard microscopy when screening for urinary tract infections in young children?** *Pediatric Infectious Disease Journal* 2001, **20:**1176-1177.
30. Benito Fernandez J, Garcia Ribes A, Trebolazabala Quirante N, Mintegui Raso S, Vazquez Ronco M, Urra Zalbidegoitia E: **[Gram stain and dipstick as diagnostic methods for urinary tract infection in febrile infants].** *Anales Espanoles de Pediatria* 2000, **53:**561-566.
31. Wammanda RD, Aikhionbare HA, Ogala WN: **Use of nitrite dipstick test in the screening for urinary tract infection in children.** *West African Journal of Medicine* 2000, **19:**206-208.
32. Bulloch B, Bausher JC, Pomerantz WJ, Connors JM, Mahabee-Gittens M, Dowd MD: **Can urine clarity exclude the diagnosis of urinary tract infection?** *Pediatrics* 2000, **106:**E60.
33. Waisman Y, Zerem E, Amir L, Mimouni M: **The validity of the uriscreen test for early detection of urinary tract infection in children.** *Pediatrics* 1999, **104:**e41.
34. Sharief N, Hameed M, Petts D: **Use of rapid dipstick tests to exclude urinary tract infection in children.** *British Journal of Biomedical Science* 1998, **55:**242-246.
35. Shaw KN, McGowan KL, Gorelick MH, Schwartz JS: **Screening for urinary tract infection in infants in the emergency department: which test is best?** *Pediatrics* 1998, **101:**E1-E5.
36. Craver RD, Abermanis JG: **Dipstick only urinalysis screen for the pediatric emergency room.** *Pediatric Nephrology* 1997, **11:**331-333.
37. Palmer LS, Richards I, Kaplan WE: **Clinical evaluation of a rapid diagnostic screen (URISCREEN) for bacteriuria in children.** *Journal of Urology* 1997, **157:**654-657.
38. Matthai J, Ramaswamy M: **Urinalysis in urinary tract infection.** *Indian Journal of Pediatrics* 1995, **62:**713-716.
39. Woodward MN, Griffiths DM: **Use of dipsticks for routine analysis of urine from children with acute abdominal pain.** *BMJ* 1993, **306:**1512.
40. Liptak GS, Campbell J, Stewart R, Hulbert WC: **Screening for urinary tract infection in children with neurogenic bladders.** *American Journal of Physical Medicine & Rehabilitation* 1993, **72:**122-126.
41. Demi M, Costa L, Zanardo V: **[Urinary tract infections in newborns: sensitivity, specificity, and predictive value of urinary screening with the reagent strip test].** *Pediatria Medica e Chirurgica* 1993, **15:**29-31.
42. Lohr JA, Portilla MG, Geuder TG, Dunn ML, Dudley SM: **Making a presumptive diagnosis of urinary tract infection by using a urinalysis performed in an on-site laboratory.** *Journal of Pediatrics* 1993, **122:**22-25.
43. Lejeune B, Baron R, Guillois B, Mayeux D: **Evaluation of a screening test for detecting urinary tract infection in newborns and infants.** *Journal of Clinical Pathology* 1991, **44:**1029-1030.
44. Tahirovic H, Pasic M: **A modified nitrite test as a screening test for significant bacteriuria.** *European Journal of Pediatrics* 1988, **147:**632-633.
45. Wiggelinkhuizen J, Maytham D, Hanslo DH: **Dipstick screening for urinary tract infection.** *South African Medical Journal* 1988, **74:**224-228.
46. Boreland PC, Stoker M: **Dipstick analysis for screening of paediatric urine.** *Journal of Clinical Pathology* 1986, **39:**1360-1362.
47. Marsik FJ, Owens D, Lewandowski J: **Use of the leukocyte esterase and nitrite tests to determine the need for culturing urine specimens from a pediatric and adolescent population.** *Diagnostic Microbiology & Infectious Disease* 1986, **4:**181-183.
48. Labbe J: **[Usefulness of testing for nitrites in the diagnosis of urinary infections in children].** *Union Medicale du Canada* 1982, **111:**261-265.
49. Fennell RS, Wilson SG, Garin EH, Pryor ND, Sorgen CD, Walker RD, Richard GA: **The combination of two screening methods in a home culture program for children with recurrent bacteriuria. An evaluation of a culture method plus a nitrite reagent test strip.** *Clinical Pediatrics* 1977, **16:**951-955.
50. Kunin CM, DeGroot JE: **Sensitivity of a nitrite indicator strip method in detecting bacteriuria in preschool girls.** *Pediatrics* 1977, **60:**244-245.
51. Todd J, McLain L, Duncan B, Brown M: **A nonculture method for home follow-up of urinary tract infections in childhood.** *Journal of Pediatrics* 1974, **85:**514-516.
52. Anad FY: **A simple method for selecting urine samples that need culturing.** *Annals of Saudi Medicine* 2001, **21:**104-105.
53. Rodriguez Cervilla J, Alonso Alonso C, Fraga Bermudez JM, Perez Munuzuri A, Gil Calvo M, Arieta Iraola G, Fernandez Lorenzo JR: **Urinary tract infection in children: Clinical and analytical prospective study for differential diagnosis in children with suspicion of an infectious disease.** *Rev Esp Pediatr* 2001, **57:**144-152.
54. Villanustre Ordonez C, Buznego Sanchez R, Rodicio Garcia M, Rodrigo Saez E, Fernandez Seara MJ, Pavon Belinchon P, Castro-Gago M: **Comparative study of semiquantitative methods (leukocytes, nitrite test and uricult) with urine culture for the diagnosis of urinary tract infection during infancy.** *An Esp Pediatr* 1994, **41:**325-328.

55. Dosa S, Houston IB, Allen LB, Jennison R: **Urinary glucose unreliable as test for urinary tract infection in infancy.** *Archives of Disease in Childhood* 1973, **48:**733-737.
56. Holland PD, Doyle CT, English L: **An evaluation of chemical tests for significant bacteriuria.** *Journal of the Irish Medical Association* 1968, **61:**128-130.
57. Kohler L, Fritz H, Schersten B: **Screening for bacteriuria with Uriglox in children.** *Acta Paediatrica Scandinavica Supplement* 1970, **206:**76-78.
58. Schersten B, Dahlqvist A, Fritz H, Kohler L, Westlund L: **Screening for bacteriuria with a test paper for glucose.** *JAMA* 1968, **204:**205-208.
59. Shaw KN, Hexter D, McGowan KL, Schwartz JS: **Clinical evaluation of a rapid screening test for urinary tract infections in children.** *Journal of Pediatrics* 1991, **118:**733-736.
60. Weinberg AG, Gan VN: **Urine screen for bacteriuria in symptomatic pediatric outpatients.** *Pediatric Infectious Disease Journal* 1991, **10:**651-654.
61. Armengol CE, Hendley JO, Schlager TA: **Urinary tract infection in young children cannot be excluded with urinalysis.** *Pediatr Res* 2000, **47:**172A.
62. Marret M, Tay S, Yap HK, Murugasu B: **Comparison of two rapid screening tests for urinary tract infection in children.** *Ann Acad Med Singap* 1995, **24:**299.
63. Parmington J, Kornberg A: **Nitrite screening for urinary tract infection in a Pediatric Emergency Department.** *Pediatr Emerg Care* 1989, **5:**285-286.
64. Giraldez M, Perozo M, González F, Rodríguez M: **Infección urinaria cinta reactiva y sedimento urinario vs. urocultivo para determinación de bacteriuria.** *Salus* 1998, **23:**27-31.
65. Lagos Zuccone R, Carter S J, Herrera Labarca P: **Utilidad de una tira reactiva y del aspecto macroscópico de la orina para descartar la sospecha clínica de infección del tracto urinario en niños ambulatorios.** *Rev Chil Pediatr* 1994, **65:**88-94.
66. Doley A, Nelligan M: **Is a negative dipstick urinalysis good enough to exclude urinary tract infection in Paediatric Emergency Department patients?** *Emergency Medicine* 2003, **15:**77-80.
67. Arslan S, Caksen H, Rastgeldi L, Uner A, Oner AF, Odabas D: **Use of urinary gram stain for detection of urinary tract infection in childhood.** *Yale Journal of Biology and Medicine* 2002, **75:**73-78.
68. Rodriguez Caballero AM, Novoa Vazquez P, Perez Ruiz A, Carmona Perez A, Cano Fernandez J, Sanchez Bayle M: **Assessment of leukocyturia in the diagnosis of urinary tract infections.** *Revista Espanola de Pediatria* 2001, **57:**305-308.
69. Hiraoka M, Hida Y, Hori C, Tsuchida S, Kuroda M, Sudo M: **Urine microscopy on a counting chamber for diagnosis of urinary infection.** *Acta Paediatrica Japonica* 1995, **37:**27-30.
70. Hoberman A, Wald ER, Reynolds EA, Penchansky L, Charron M: **Is urine culture necessary to rule out urinary tract infection in young febrile children?** *Pediatric Infectious Disease Journal* 1996, **15:**304-309.
71. Hoberman A, Wald ER, Reynolds EA, Penchansky L, Charron M: **Pyuria and bacteriuria in urine specimens obtained by catheter from young children with fever.** *Journal of Pediatrics* 1994, **124:**513-519.
72. Lin DS, Huang SH, Lin CC, Tung YC, Huang TT, Chiu NC, Koa HA, Hung HY, Hsu CH, Hsieh WS, Yang DI, Huang FY: **Urinary tract infection in febrile infants younger than eight weeks of age.** *Pediatrics* 2000, **105:**E20-20.
73. Lin DS, Huang FY, Chiu NC, Koa HA, Hung HY, Hsu CH, Hsieh WS, Yang DI: **Comparison of hemocytometer leukocyte counts and standard urinalyses for predicting urinary tract infections in febrile infants.** *Pediatric Infectious Disease Journal* 2000, **19:**223-227.
74. Pryles CV, Eliot CR: **Pyuria and bacteriuria in infants and children. The value of pyuria as a diagnostic criterion of urinary tract infections.** *American Journal of Diseases of Children* 1965, **110:**628-635.
75. Santos MA, Mos EN, Schmidt BJ, Piva S: **Comparacion entre el estudio bacterioscopico cuantitativo y el urocultivo para el diagnostico de infeccion urinaria en pediatria.** *Bol Méd Hosp Infant Méx* 1982, **39:**526-530.
76. Saxena H, Ajwani KD, Mehrotra D: **Quantitative pyuria in the diagnosis of urinary infections in children.** *Indian Journal of Pediatrics* 1975, **42:**35-38.
77. Schreiter G, Buhtz P: **[Diagnostic value of the cytologic and bacteriologic urine examinations in pediatrics. II. Comparison of leukocyturia and bacteriuria].** *Deutsche Gesundheitswesen* 1971, **26:**1318-1323.
78. Littlewood JM, Jacobs SI, Ramsden CH: **Comparison between microscopical examination of unstained deposits of urine and quantitative culture.** *Archives of Disease in Childhood* 1977, **52:**894-896.
79. Lockhart GR, Lewander WJ, Cimini DM, Josephson SL, Linakis JG: **Use of urinary gram stain for detection of urinary tract infection in infants.** *Annals of Emergency Medicine* 1995, **25:**31-35.
80. Purwar VN, Agrawal SP, Dikshit SK: **Gram stained urine slides in the diagnosis of urinary tract infections in children.** *Journal of the Indian Medical Association* 1972, **59:**387-388.
81. Vangone G, Russo G: **[Bacteria and leukocyte count in the urine in the diagnosis of urinary tract infections].** *Pediatria Medica e Chirurgica* 1985, **7:**125-129.
82. Vickers D, Ahmad T, Coulthard MG: **Diagnosis of urinary tract infection in children: fresh urine microscopy or culture?** *Lancet* 1991, **338:**767-770.
83. Manson R, Scholefield J, Johnston RJ, Scott R: **The screening of more than 2,000 schoolgirls for bacteriuria using an automated fluorescence microscopy system.** *Urological Research* 1985, **13:**143-148.
84. Hoberman A, Wald ER, Penchansky L, Reynolds EA, Young S: **Enhanced urinalysis as a screening test for urinary tract infection.** *Pediatrics* 1993, **91:**1196-1199.
85. Bachur R, Harper MB: **Reliability of the urinalysis for predicting urinary tract infections in young febrile children.** *Archives of Pediatrics & Adolescent Medicine* 2001, **155:**60-65.
86. Dickersin K, Chan S, Chalmers TC, Sacks HS, Smith H: **Publication bias and clinical trials.** *Control Clin Trials* 1987, **8:**343-353.
87. Begg CB, Berlin JA: **Publication bias and dissemination of clinical research.** *J Natl Cancer Inst* 1989, **81:**107-115.
88. Steurer J, Fischer JE, Bachmann LM, Koller M, ter Riet G: **Communicating test accuracy terms to practicing physicians: a controlled study.** *BMJ* 2002, **324:**824-826.
89. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PMM, Zwinderman AH: **Direct pooling of sensitivity and specificity using bivariate models in meta-analysis of studies of diagnostic accuracy.** *Submitted* .
90. Bossuyt PMM, Reitsma JB, Bruns DE, Gatsonis C, Glasziou P, Irwig L, Moher D, Rennie D, de Vet HCW, Lijmer JG: **The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration.** *Annals of Internal Medicine* 2003, **138:**W1-W12.

Pre-publication history

The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1471-2431/5/4/prepub